

# THE BOUNDEDNESS-BY-ENTROPY PRINCIPLE FOR CROSS-DIFFUSION SYSTEMS

ANSGAR JÜNGEL

**ABSTRACT.** A novel principle is presented which allows for the proof of bounded weak solutions to a class of physically relevant, strongly coupled parabolic systems exhibiting a formal gradient-flow structure. The main feature of these systems is that the diffusion matrix may be generally neither symmetric nor positive semi-definite. The key idea of the principle is to employ a transformation of variables, determined by the entropy density, which is defined by the gradient-flow formulation. The transformation yields at the same time a positive semi-definite diffusion matrix, suitable gradient estimates as well as lower and/or upper bounds of the solutions. These bounds are a consequence of the transformation of variables and are obtained without the use of a maximum principle. Several classes of cross-diffusion systems are identified which can be solved by this technique. The systems are formally derived from continuous-time random walks on a lattice modeling, for instance, the motion of ions, cells, or fluid particles.

## CONTENTS

1. Introduction	2
1.1. Idea of the principle	2
1.2. An illustrative example	3
1.3. State of the art	5
1.4. Main results	5
2. Examples	9
2.1. Examples with volume filling	10
2.2. Models without volume filling	11
3. Proof of Theorem 2	12
4. Proof of Theorem 3	17
5. Proof of Theorem 4	24
6. Further results and open problems	28
Appendix A. Relations to non-equilibrium thermodynamics	32
Appendix B. Formal derivation of two-species population models	33
Appendix C. A variant of the Aubin compactness lemma	34
References	36

---

*Date:* March 24, 2014.

*2000 Mathematics Subject Classification.* 35K51, 35K65, 35Q92, 92D25.

*Key words and phrases.* Strongly coupled parabolic systems, global-in-time existence, bounded weak solutions, gradient flow, entropy variables, entropy method, volume-filling effects, population dynamics models, non-equilibrium thermodynamics.

The author acknowledges partial support from the Austrian Science Fund (FWF), grants P22108, P24304, and W1245, and the Austrian-French Program of the Austrian Exchange Service (ÖAD). Part of this manuscript was written during the stay of the author at the King Abdullah University of Science and Technology (KAUST) in Thuwal, Saudi-Arabia. The author thanks Peter Markowich for his kind invitation and support.

## 1. INTRODUCTION

Many applications in physics, chemistry, and biology can be modeled by reaction-diffusion systems with cross diffusion, which describe the temporal evolution of the densities or mass fractions of a multicomponent system. Physically, we expect that the concentrations are nonnegative or even bounded (examples are given in Section 2). Since generally no maximum principle holds for parabolic systems, the proof of these bounds is a challenging problem. A second difficulty arises from the fact that in many applications, the diffusion matrix is neither symmetric nor positive semi-definite.

In this paper, we introduce a general technique which allows us, under certain assumptions, to prove simultaneously the global existence of a weak solution as well as its boundedness from below and/or above. The key idea is to exploit the so-called entropy structure of the parabolic system, which is assumed to exist, leading at the same time to gradient estimates and lower/upper bounds. More specifically, we consider reaction-diffusion systems of the form

$$(1) \quad \partial_t u - \operatorname{div}(A(u)\nabla u) = f(u) \quad \text{in } \Omega, \quad t > 0,$$

subject to the boundary and initial conditions

$$(2) \quad (A(u)\nabla u) \cdot \nu = 0 \quad \text{on } \partial\Omega, \quad t > 0, \quad u(0) = u^0 \quad \text{in } \Omega.$$

Here,  $u(t) = (u_1, \dots, u_n)(\cdot, t)$  is a vector-valued function ( $n \geq 1$ ), representing the densities or mass fractions  $u_i$  of the components of the system,  $A(u) = (a_{ij}(u)) \in \mathbb{R}^{n \times n}$  is the diffusion matrix, and the reactions are modeled by the components of the function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Furthermore,  $\Omega \subset \mathbb{R}^d$  ( $d \geq 1$ ) is a bounded domain with Lipschitz boundary and  $\nu$  is the exterior unit normal vector to  $\partial\Omega$ . The divergence  $\operatorname{div}(A(u)\nabla u)$  and the expression  $(A(u)\nabla u) \cdot \nu$  consist of the components

$$\sum_{j=1}^d \sum_{k=1}^n \frac{\partial}{\partial x_j} \left( a_{ik}(u) \frac{\partial u_k}{\partial x_j} \right), \quad \sum_{j=1}^d \sum_{k=1}^n a_{ik}(u) \frac{\partial u_k}{\partial x_j} \nu_j, \quad i = 1, \dots, n,$$

respectively. From the applications, we expect that the solution  $u$  has values in an open set  $D \subset \mathbb{R}^n$ . When  $u$  models concentrations, we expect that  $D \subset (0, \infty)^n$  (positivity); when  $u$  models mass fractions, we expect that the values of each component  $u_i$  are between zero and one, i.e.  $D \subset (0, 1)^n$  (boundedness and positivity).

Before summarizing the state of the art of cross-diffusion systems and detailing our main results, we explain the key idea of our principle and provide an illustrating example.

**1.1. Idea of the principle.** The main assumption in this paper is that system (1) possesses a formal gradient-flow structure, i.e., (1) can be formulated as

$$\partial_t u - \operatorname{div} \left( B \nabla \frac{\delta \mathcal{H}}{\delta u} \right) = f(u),$$

where  $B$  is a positive semi-definite matrix and  $\delta \mathcal{H} / \delta u$  is the variational derivative of the entropy (or free energy) functional  $\mathcal{H}[u] = \int_{\Omega} h(u) dx$ . The function  $h : D \rightarrow [0, \infty)$  is called the entropy density. Identifying  $\delta \mathcal{H} / \delta u$  with its Riesz representative  $Dh(u)$  (the derivative of  $h$ ) and introducing the entropy variable  $w = Dh(u)$ , the above formulation can be understood as

$$(3) \quad \partial_t u - \operatorname{div}(B(w)\nabla w) = f(u),$$

where  $B = B(w) = A(u)(D^2h(u))^{-1}$  and  $D^2h$  is the Hessian of  $h$ . For transforming back from the  $w$ - to the  $u$ -variable, we need to assume that  $Dh : D \rightarrow \mathbb{R}^n$  is invertible such that  $u = (Dh)^{-1}(w)$ .

The gradient-flow formulation has two important consequences. First, calculating the formal time derivative of  $\mathcal{H}$ , (3) and integrating by parts yields

$$(4) \quad \frac{d\mathcal{H}}{dt} = \int_{\Omega} \partial_t u \cdot Dh(u) dx = \int_{\Omega} \partial_t u \cdot w dx = - \int_{\Omega} \nabla w : B(w) \nabla w dx + \int_{\Omega} f(u) \cdot w dx,$$

where  $A : B = \sum_{i,j} a_{ij} b_{ij}$  for matrices  $A = (a_{ij})$  and  $B = (b_{ij})$ . Thus, if  $f(u) \cdot w \leq 0$  and since  $B(w)$  is assumed to be positive semi-definite,  $\mathcal{H}$  is a Lyapunov functional. We refer to  $\mathcal{H}$  as an entropy and to the integral of  $\nabla w : B \nabla w$  as the corresponding entropy dissipation. Under certain conditions, it gives gradient estimates for  $u$  needed to prove the global-in-time existence of solutions to (2) and (3).

We remark that the positive semi-definiteness of  $B(w)$  is in fact a consequence of the existence of an entropy. It was shown in [19, 35] that both properties are equivalent and moreover,  $B(w)$  may be even symmetric.

Second, supposing that there exists a weak solution  $w$  to (3), the invertibility of  $Dh : D \rightarrow \mathbb{R}^n$  shows that the original variable  $u = (Dh)^{-1}(w)$  satisfies  $u(\cdot, t) \in D$  for  $t > 0$ . Thus, if  $D$  is bounded, we automatically obtain  $L^\infty$  bounds for  $u$ , without using a maximum principle. If  $D$  is only a cone, for instance  $D = (0, \infty)^n$ , we conclude the positivity of  $u(t)$ .

We call the above technique the *boundedness-by-entropy principle* since it provides lower and/or upper bounds for the solutions to (1)-(2) by the use of the entropy density. Summarizing, the principle is based on two main hypotheses:

H1: There exists a function  $h \in C^2(D; [0, \infty))$  whose derivative is invertible on  $\mathbb{R}^n$ . This yields the bound  $u(\cdot, t) \in D$ .

H2: The matrix  $D^2h(u)A(u)$  is positive semi-definite for all  $u \in D$ . This condition is necessary to derive a priori estimates for  $u$ .

Note that the positive semi-definiteness of  $D^2h(u)A(u)$  is equivalent to that of  $B(w) = A(u)(D^2h(u))^{-1}$  since for all  $z \in \mathbb{R}^n$ ,  $z^\top D^2h(u)A(u)z = (D^2h(u)z)^\top B(w)(D^2h(u)z)$ . Hypothesis H2 avoids the inversion of  $D^2h(u)$ .

In fact, we need a stronger hypothesis than H2 since it does not allow us to infer gradient estimates. We need to suppose that  $D^2h(u)A(u)$  is positive definite in such a way that we obtain  $L^2$  gradient estimates for  $u_i^m$ , where  $m > 0$  is some number. Moreover, we need an estimate for the time derivative of  $u_i$  which makes it necessary to impose some growth conditions on the coefficients of  $A(u)$ . We explain these requirements with the help of the following example.

**1.2. An illustrative example.** We consider a multicomponent fluid consisting of three components with mass fractions  $u_1$ ,  $u_2$ , and  $1 - u_1 - u_2$  and equal molar masses under isobaric, isothermal conditions. The model equals (1) with the diffusion matrix

$$(5) \quad A(u) = \frac{1}{2 + 4u_1 + u_2} \begin{pmatrix} 1 + 2u_1 & u_1 \\ 2u_2 & 2 + u_2 \end{pmatrix}$$

where we have chosen particular diffusivities to simplify the presentation (see Section 2.1). Notice that the nonnegativity of  $u_1$  and  $u_2$  can be proved easily by a maximum principle argument but the proof of upper bounds is less clear. The logarithmic entropy density

$$(6) \quad h(u) = u_1(\log u_1 - 1) + u_2(\log u_2 - 1) + (1 - u_1 - u_2)(\log(1 - u_1 - u_2) - 1)$$

for  $u = (u_1, u_2) \in D$ , where

$$(7) \quad D = \{(u_1, u_2) \in (0, 1)^2 : u_1 + u_2 < 1\}$$

satisfies Hypothesis H1, since the inverse transformation of variables gives

$$(8) \quad u = (Dh)^{-1}(w) = \left( \frac{e^{w_1}}{1 + e^{w_1} + e^{w_2}}, \frac{e^{w_2}}{1 + e^{w_1} + e^{w_2}} \right) \in D,$$

where  $w = (w_1, w_2) \in \mathbb{R}^2$ . Thus, once the existence of weak solutions  $w$  to the transformed system (3) is shown, we conclude the bounds  $0 < u_1, u_2 < 1$  *automatically* by transforming back to the original variable. In particular, no maximum principle is used. In fact, the inverse transformation even shows that  $1 - u_1 - u_2 > 0$ , as required from the application. Furthermore, the matrix

$$D^2h(u)A(u) = \frac{1}{\delta(u)} \begin{pmatrix} u_2(1 - u_1 - u_2) + 3u_1u_2 & 3u_1u_2 \\ 3u_1u_2 & 2u_1(1 - u_1 - u_2) + 3u_1u_2 \end{pmatrix}$$

with  $\delta(u) = u_1u_2(1 - u_1 - u_2)(2 + 4u_1 + u_2)$  is symmetric and positive definite, thus fulfilling Hypothesis H2.

However, the matrix  $B = A(u)(D^2h(u))^{-1}$  degenerates at  $u_1 = 0$  or  $u_2 = 0$ , and we cannot expect to conclude gradient estimates for  $w$  from (4). Instead, it is more appropriate to derive these estimates for the original variable  $u$ . Indeed, calculating the time derivative of the entropy according to (4) and using  $\nabla w = D^2h(u)\nabla u$ , we find that

$$\frac{d\mathcal{H}}{dt} + \int_{\Omega} \nabla u : D^2h(u)A(u)\nabla u dx = \int_{\Omega} f(u) \cdot w dx,$$

and the entropy dissipation can be estimated according to

$$\begin{aligned} \int_{\Omega} \nabla u : D^2h(u)A(u)\nabla u dx &= \int_{\Omega} \frac{1}{2 + 4u_1 + u_2} \left( \frac{|\nabla u_1|^2}{u_1} + \frac{2|\nabla u_2|^2}{u_2} + \frac{3|\nabla(u_1 + u_2)|^2}{1 - u_1 - u_2} \right) dx \\ &\geq \int_{\Omega} (2|\nabla \sqrt{u_1}|^2 + 4|\nabla \sqrt{u_2}|^2) dx. \end{aligned}$$

Thus, assuming that the integral involving the reaction terms can be bounded uniformly in  $u$ , we obtain  $H^1(\Omega)$  estimates for  $u_1^{1/2}$  and  $u_2^{1/2}$ . Using the boundedness of  $u_i$ , a priori estimates for  $\partial_t u_1$  and  $\partial_t u_2$  can be proven, taking into account the particular structure of  $A(u)$ .

This example shows that we need additional assumptions on the nonlinearities of system (1) in order to derive suitable a priori estimates, detailed in Section 1.4.

**Remark 1** (Notion of entropy). There exists an intimate relation between the boundedness-by-entropy principle and non-equilibrium thermodynamics. In particular, the entropy variable  $w = Dh(u)$  is related to the chemical potentials of a mixture of gases and the special transformation (8) is connected with a special choice of thermodynamic activities; see Appendix A for details. The entropy density defined above equals the negative thermodynamic entropy. Since the physical entropy is increasing and we wish to investigate nonincreasing functionals, we have reversed the sign as usual in entropy methods. Moreover, the logarithmic entropy (6) is motivated by Boltzmann's entropy for kinetic equations. For these reasons, we refer to the functional  $\mathcal{H}[u] = \int_{\Omega} h(u) dx$  as a (mathematical) entropy. We note that in some applications, free energy may be a more appropriate notion.  $\square$

**1.3. State of the art.** We have already mentioned that the analysis of cross-diffusion systems is delicate since standard tools like maximum principles and regularity results generally do not apply. For instance, there exist Hölder continuous solutions to certain cross-diffusion systems which are not bounded, and there exist bounded weak solutions which develop singularities in finite time [56]. In view of these counterexamples, it is not surprising that additional conditions are required to prove that (local in time) weak solutions are bounded and that they can be continued globally in time.

Ladyzenskaya et al. [38, Chap. VII] reduce the problem of finding a priori estimates of local-in-time solutions  $u$  to quasilinear parabolic systems to the problem of deriving  $L^\infty$  bounds for  $u$  and  $\nabla u$ . Under some growth conditions on the nonlinearities, the global-in-time existence of classical solutions was shown. A fundamental theory of strongly coupled systems was developed by Amann [2]. He formulated the concept of  $W^{1,p}$  weak solutions and their local existence and proved that the solutions exist globally if their  $L^\infty$  and Hölder norms can be controlled. The above mentioned counterexamples show that the control on both norms is necessary. Le and Nguyen [41] proved that bounded weak solutions are Hölder continuous if certain structural assumptions on the diffusion matrix are imposed. The regularity of the solutions to systems with diagonal or full diffusion matrix was investigated in, for instance, [3, 26, 44, 50].

The boundedness of weak solutions to strongly coupled systems has been proved using various methods. Invariance principles were employed by Küfner [37] and Redlinger [51], requiring severe restrictions on the initial data. Truncated test functions, which are nonlinear in the solutions to  $2 \times 2$  systems, were suggested by Le [39] who proved the boundedness under rather restrictive structural assumptions. In the work of Lepoutre et al. [42], the existence of bounded solutions to strongly coupled systems with spatially regularized arguments in the nonlinearities was shown using Hölder theory for nondivergence parabolic operators. Other methods are based on the derivation of  $L^p$  bounds uniform in  $p$  and the passage to the limit  $p \rightarrow \infty$  (Moser-type or Alikakos-type iterations [1]).

The idea of proving the boundedness of weak solutions using the entropy density (6) was, to our best knowledge, first employed by Burger et al. [8] in a size-exclusion model for two species (see Section 2.1). It was applied to a tumor-growth model in [33] and extended to Maxwell-Stefan systems for fluids with arbitrary many components [34]. A different entropy density was suggested in [29] to prove  $L^\infty$  bounds in one space dimension. In fact, the idea of using entropy variables already appeared in the analysis of parabolic systems from non-equilibrium thermodynamics [19], originally used to remove the influence from the electric potential, and goes back to the use of so-called Slotboom variables in semiconductor modeling [46, Section 3.2].

In this paper, we identify the key elements of this idea and provide a general global existence result for bounded weak solutions to certain systems. Furthermore, the technique is applied to a variety of cross-diffusion systems derived from a random-walk master equation on a lattice, underlying the strength and flexibility of the method.

**1.4. Main results.** Our first main result concerns the existence of bounded weak solutions to (1)-(2) under some general structural assumptions. Motivated by the comments in Sections 1.1 and 1.2, we impose the following hypotheses:

- H1: There exists a convex function  $h \in C^2(D; [0, \infty))$  ( $D \subset \mathbb{R}^n$  open,  $n \geq 1$ ) such that its derivative  $Dh : D \rightarrow \mathbb{R}^n$  is invertible on  $\mathbb{R}^n$ .

H2': Let  $D \subset (a, b)^n$  for some  $a, b \in \mathbb{R}$  with  $a < b$  and let  $\alpha_i^*, m_i \geq 0$  ( $i = 1, \dots, n$ ) be such that for all  $z = (z_1, \dots, z_n)^\top \in \mathbb{R}^n$  and  $u = (u_1, \dots, u_n) \in D$ ,

$$z^\top D^2 h(u) A(u) z \geq \sum_{i=1}^n \alpha_i(u_i)^2 z_i^2,$$

where either  $\alpha_i(u_i) = \alpha_i^*(u_i - a)^{m_i-1}$  or  $\alpha_i(u_i) = \alpha_i^*(b - u_i)^{m_i-1}$ .

H2'': There exists  $a^* > 0$  such that for all  $u \in D$  and  $i, j = 1, \dots, n$  for which  $m_j > 1$ , it holds that  $|a_{ij}(u)| \leq a^* |\alpha_j(u_j)|$ .

H3: It holds  $A \in C^0(D; \mathbb{R}^{n \times n})$  and there exists  $C_f > 0$  such that for all  $u \in D$ ,  $f(u) \cdot Dh(u) \leq C_f(1 + h(u))$ .

Hypothesis H1 shows that the inverse transformation  $u = (Dh)^{-1}(w)$  is well defined. If  $w(t)$  is a weak solution to (3), we conclude that  $u(t) = (Dh)^{-1}(w(t)) \in D$ , yielding the desired  $L^\infty$  bounds on  $u(t)$ . Assumption H2', which implies H2, is employed to prove a priori estimates for  $\nabla u_i^{m_i}$ . Hypothesis H2'' is needed to show a bound on  $\partial_t u_i$ . The condition on  $f(u)$  in Hypothesis H3 is needed to derive an a priori estimate for the solution. For instance, if the entropy density is given by (6), we may choose  $f_i(u) = u_i^{\alpha_i} (1 - u_1 - u_2)^{\beta_i}$  for  $\alpha_i, \beta_i > 0$  and  $i = 1, 2$ . It is related to the quasi-positivity assumption  $f_i(u) \geq 0$  for all  $u \in (0, \infty)^2$  with  $u_i = 0$  [4, Section 6]. In contrast to the structural assumptions of [39], Hypotheses H1-H3 are easy to verify as soon as an entropy density is found (often motivated from the application at hand). Under the above conditions, the following theorem holds.

**Theorem 2** (General global existence result). *Let Hypotheses H1, H2', H2'', and H3 hold and let  $u^0 \in L^1(\Omega; \mathbb{R}^n)$  be such that  $u^0(x) \in D$  for  $x \in \Omega$ . Then there exists a weak solution  $u$  to (1)-(2) satisfying  $u(x, t) \in \overline{D}$  for  $x \in \Omega$ ,  $t > 0$  and*

$$u \in L_{\text{loc}}^2(0, \infty; H^1(\Omega; \mathbb{R}^n)), \quad \partial_t u \in L_{\text{loc}}^2(0, \infty; H^1(\Omega; \mathbb{R}^n)').$$

*The initial datum is satisfied in the sense of  $H^1(\Omega; \mathbb{R}^n)'$ .*

Note that since  $D$  is bounded by Hypothesis H2', the theorem yields global  $L^\infty$  bounds on  $u$ . We remark that we may also assume the slightly weaker condition  $u^0(x) \in \overline{D}$ . Indeed, we may approximate  $u^0$  by  $u_\eta^0(x) \in D$  satisfying  $u_\eta^0 \rightarrow u^0$  a.e., apply the theorem to  $u_\eta^0$ , and perform the limit  $\eta \rightarrow 0$ . We refer to [17] for details.

The assumptions of the theorem are satisfied for the tumor-growth model and the Maxwell-Stefan equations – detailed in Section 2.1 –, including the example of Section 1.2. Furthermore, the diffusion matrix

$$(9) \quad A(u) = \begin{pmatrix} 1 - u_1 & -u_1 \\ -u_2 & 1 - u_2 \end{pmatrix}$$

with entropy density (6) and domain (7) also satisfies Hypotheses H1, H2', and H2''. The model with this diffusion matrix describes the aggregation of two population species with cross-diffusion terms related to the drift term of the Keller-Segel system.

For the proof of Theorem 2, we first semi-discretize (3) in time with step size  $\tau > 0$  and regularize this equation by the expression  $\varepsilon(\sum_{|\alpha|=m} (-1)^m D^{2\alpha} w + w)$ , where  $\varepsilon > 0$ ,  $D^{2\alpha}$  is a partial derivative of order  $2|\alpha|$ , and  $m \in \mathbb{N}$  is such that  $H^m(\Omega) \hookrightarrow L^\infty(\Omega)$ . This regularization ensures that the approximate solution  $w^{(\tau)}$  is bounded. For the proof of approximate solutions, we only need Hypotheses H1, H2, and H3. The discrete version of the entropy identity (4) and Hypothesis H2' yield gradient estimates for  $u^{(\tau)} = (Dh)^{-1}(w^{(\tau)})$ . Furthermore, uniform

estimates on the discrete time derivative of  $u^{(\tau)}$  can be shown with the help of Hypothesis H2". Then, by the discrete Aubin lemma from [25], the limit  $(\tau, \varepsilon) \rightarrow 0$  can be performed yielding the existence of a weak solution to (1)-(2).

Assumptions H2' and H2" may be too restrictive in certain applications. For instance, it may be impossible to find a bounded set  $D \subset \mathbb{R}^n$  satisfying H1 or the inequality in H2' is not satisfied (see the examples in Section 2). However, we show that our technique can be adapted to situations in which *variants* of Hypotheses H2' and H2" hold. To illustrate this idea, we choose a class of cross-diffusion systems modeling the time evolution of two population species. These models are derived from a random-walk master equation in the diffusive limit (see Appendix B). They generalize several models from the literature (see Section 2).

We consider two situations. In the first case, we assume that volume limitations lead to a limitation of the population densities (volume-filling case). Then the densities are nonnegative and bounded by a threshold value which is normalized to one. This situation occurs, for instance, in the volume-filling Keller-Segel model [48] and in ion-channel modeling [30]. The diffusion matrix in (1) reads as (see Appendix B)

$$(10) \quad A(u) = \begin{pmatrix} q(u_3) + u_1 q'(u_3) & u_1 q'(u_3) \\ \beta u_2 q'(u_3) & \beta(q(u_3) + u_2 q'(u_3)) \end{pmatrix},$$

where  $u_1$  and  $u_2$  are the densities of the species with bounded total density,  $u_1 + u_2 \leq 1$ , and  $u_3 = 1 - u_1 - u_2$ . The function  $q$  is related to the transition probability of a species to move from one cell to a neighboring cell, and  $\beta > 0$  is the ratio between the transition rates of both species. Biologically,  $q$  vanishes when the cells are fully packed, i.e. if  $u_1 + u_2 = 1$ , so  $q(0) = 0$  and  $q$  is nondecreasing. When only one species is considered, the diffusion equation corresponds to the equation for the cell density in the volume-filling chemotaxis model [60]. The special case  $q(u_3) = u_3$  was analyzed by Burger et al. [8]. We are able to prove the global existence of bounded weak solutions for  $q(u_3) = u_3^s$  for any  $s \geq 1$  but also more general functions are allowed (see Theorems 6 and 9).

**Theorem 3** (Volume-filling case). *Let  $s \geq 1$ ,  $\beta > 0$ , and  $q(y) = y^s$  for  $y \geq 0$ . Furthermore, let  $u^0 = (u_1^0, u_2^0) \in L^1(\Omega; \mathbb{R}^2)$  with  $u_1^0, u_2^0 \geq 0$ ,  $u_1^0 + u_2^0 \leq 1$  in  $\Omega$ . Then there exists a bounded weak solution  $u = (u_1, u_2)$  to (1)-(2) with diffusion matrix (10) and  $f = 0$  satisfying  $0 \leq u_1, u_2 \leq 1$  and  $u_3 := 1 - u_1 - u_2 \geq 0$  in  $\Omega$ ,  $t > 0$ ,*

$$u_i q(u_3)^{1/2}, q(u_3)^{1/2} \in L_{\text{loc}}^2(0, \infty; H^1(\Omega; \mathbb{R}^2)), \quad \partial_t u_i \in L_{\text{loc}}^2(0, \infty; H^1(\Omega; \mathbb{R}^2)')$$

for  $i = 1, 2$ , and for all  $T > 0$  and  $\phi = (\phi_1, \phi_2) \in L^2(0, T; H^1(\Omega))^2$ ,

$$(11) \quad \int_0^T \langle \partial_t u, \phi \rangle dt + \sum_{i=1}^2 \beta_i \int_0^T \int_{\Omega} (q(u_3)^{1/2} \nabla(q(u_3)^{1/2} u_i) - 3q(u_3)^{1/2} u_i \nabla(q(u_3)^{1/2})) \cdot \nabla \phi_i dx dt = 0,$$

where  $\beta_1 = 1$ ,  $\beta_2 = \beta$  and  $\langle \cdot, \cdot \rangle$  is the dual product between  $H^1(\Omega; \mathbb{R}^2)'$  and  $H^1(\Omega; \mathbb{R}^2)$ . The initial datum is satisfied in the sense of  $H^1(\Omega; \mathbb{R}^2)'$ .

We may also consider reaction terms  $f(u)$  satisfying a particular structure; see Remark 8. The key idea of the proof is to introduce the entropy density

$$(12) \quad h(u) = u_1(\log u_1 - 1) + u_2(\log u_2 - 1) + \int_c^{u_3} \log q(y) dy,$$

where  $u = (u_1, u_2) \in D = \{(u_1, u_2) \in (0, 1)^2 : u_1 + u_2 < 1\}$  and  $0 < c < 1$ . A computation shows that Hypotheses H1 and H2 are satisfied but not H2'. Indeed, we will prove in Section 4 (see (32)) that

$$\nabla u^\top D^2 h(u) A(u) \nabla u \geq \frac{q(u_3)}{u_1} |\nabla u_1|^2 + \frac{q(u_3)}{u_2} |\nabla u_2|^2,$$

and the factors degenerate at  $u_1 + u_2 = 1$  since  $q(0) = 0$ . From this, we are able to conclude a gradient estimate for  $u_3$  but not for  $u_1$  or  $u_2$ . However, this is sufficient to infer the existence of solutions, employing an extension of Aubin's compactness lemma (see Appendix C).

In the second case, volume-filling effects are not taken into account. Then the diffusion matrix reads as (see Appendix B)

$$(13) \quad A(u) = \begin{pmatrix} p_1(u) + u_1 \partial_1 p_1(u) & u_1 \partial_2 p_1(u) \\ u_2 \partial_1 p_2(u) & p_2(u) + u_2 \partial_2 p_2(u) \end{pmatrix},$$

where  $p_1$  and  $p_2$  are related to the transition probabilities of the two species and  $\partial_i p_j = \partial p_j / \partial u_i$ . Note that each row of  $A(u)$  is the gradient of a function such that

$$\operatorname{div}(A(u) \nabla u)_i = \Delta(u_i p_i(u)), \quad i = 1, 2,$$

which allows for additional  $L^2$  estimates for  $u_1$  and  $u_2$ , using the duality estimates of Pierre and Schmitt [49]. This observation was exploited in [23].

When the transition probabilities depend linearly on the densities,  $p_i(u) = \alpha_{i0} + \alpha_{i1} u_1 + \alpha_{i2} u_2$  ( $i = 1, 2$ ), we obtain the well-known population model of Shigesada, Kawashima, and Teramoto [54] with the diffusion matrix

$$(14) \quad A(u) = \begin{pmatrix} \alpha_{01} + 2\alpha_{11}u_1 + \alpha_{12}u_2 & \alpha_{12}u_1 \\ \alpha_{21}u_2 & \alpha_{20} + \alpha_{21}u_1 + 2\alpha_{22}u_2 \end{pmatrix}.$$

The maximum principle implies that  $u_1$  and  $u_2$  are nonnegative. Less results are known concerning upper bounds. In fact, in one space dimension and with coefficients  $\alpha_{10} = \alpha_{20}$ , Shim [55] proved uniform upper bounds. Moreover, if cross-diffusion is weaker than self-diffusion (i.e.  $\alpha_{12} < \alpha_{22}$ ,  $\alpha_{21} < \alpha_{11}$ ), weak solutions are bounded and Hölder continuous [39]. Upper bounds without any restrictions are not known so far (at least to our best knowledge).

This model has received a lot of attention in the mathematical literature. One of the first existence results is due to Kim [36] who neglected self-diffusion ( $\alpha_{11} = \alpha_{22} = 0$ ) and assumed equal coefficients ( $\alpha_{ij} = 1$ ). The tridiagonal case  $\alpha_{21} = 0$  was investigated by Amann [2], Le [40], and more recently, by Desvillettes and Trescases [24]. Yagi [61] proved an existence theorem under the assumption that the diffusion matrix is positive definite ( $\alpha_{12} < 8\alpha_{11}$ ,  $\alpha_{21} < 8\alpha_{22}$ ,  $\alpha_{12} = \alpha_{21}$ ). The first global existence result without any restriction on the diffusion coefficients (except positivity) was achieved in [29] in one space dimension and in [14, 15] in several space dimensions. The case of concave functions  $p_1$  and  $p_2$ , for instance,

$$(15) \quad p_i(u) = \alpha_{i0} + a_{i1}(u_1) + a_{i2}(u_2), \quad \text{where } a_{i1}(u_1) = \alpha_{i1}u_1^s, \quad a_{i2}(u_2) = \alpha_{i2}u_2^s,$$

and  $i = 1, 2$ ,  $u = (u_1, u_2)$ ,  $0 < s < 1$ , was analyzed by Desvillettes et al. [23]. We are able to generalize this result to the case  $1 < s < 4$  but we need to restrict the size of the cross-diffusion coefficients  $\alpha_{12}$  and  $\alpha_{21}$ . More general functions  $p_1$  and  $p_2$  are possible; see Section 5.

**Theorem 4** (No volume-filling case). *Let (15) hold with  $1 < s < 4$  and  $(1 - 1/s)\alpha_{12}\alpha_{21} \leq \alpha_{11}\alpha_{22}$ . Furthermore, let  $u^0 = (u_1^0, u_2^0) \in L^1(\Omega; \mathbb{R}^2)$  with  $u_1^0, u_2^0 \geq 0$  in  $\Omega$  and  $\int_\Omega h(u^0)dx < \infty$ ,*



where  $h$  is defined in (16) below. Then there exists a weak solution  $u = (u_1, u_2)$  to (1)-(2) with diffusion matrix (13) and  $f = 0$  satisfying  $u_i \geq 0$  in  $\Omega$ ,  $t > 0$ ,

$$u_i^{s/2}, u_i^s \in L_{\text{loc}}^2(0, \infty; H^1(\Omega; \mathbb{R}^2)), \quad u_i \in L_{\text{loc}}^\infty(0, \infty; L^s(\Omega; \mathbb{R}^2)), \\ \partial_t u_i \in L_{\text{loc}}^1(0, \infty; X'), \quad i = 1, 2,$$

where  $X = \{\psi \in W^{m, \infty}(\Omega) : \nabla \psi \cdot \nu = 0 \text{ on } \partial\Omega\}$ ,  $m > d/2$ , for all  $\phi = (\phi_1, \phi_2) \in L^\infty(0, T; X)^2$ ,

$$\int_0^T \langle \partial_t u, \phi \rangle dt + \sum_{i=1}^2 \int_0^T \int_\Omega u_i p_i(u) \Delta \phi dx dt = 0,$$

and  $u(0) = u^0$  in the sense of  $X'$ .

Nonvanishing reaction terms  $f(u)$  can be treated if they satisfy an appropriate growth condition such that  $f(u^{(\tau)})$  is bounded in some  $L^p$  space with  $p > 1$ , where  $u^{(\tau)}$  is a solution to an approximate problem. We leave the details to the reader.

The idea of the proof is to employ the entropy density

$$(16) \quad h(u) = \int_c^{u_1} \int_c^z \frac{a'_{21}(y)}{y} dy dz + \int_c^{u_2} \int_c^z \frac{a'_{12}(y)}{y} dy dz$$

for  $u = (u_1, u_2) \in D = (0, \infty)^2$ , where  $c > 0$ . Hypothesis H2 is only satisfied if the restriction  $(1 - 1/s)\alpha_{12}\alpha_{21} \leq \alpha_{11}\alpha_{22}$  holds. It is an open problem whether there exists another entropy density fulfilling Hypothesis H2 without any restriction on  $\alpha_{ij}$  (except positivity). In order to satisfy Hypothesis H1, we need to regularize the entropy density (16) as in [23]:

$$h_\varepsilon(u) = h(u) + \varepsilon u_1(\log u_1 - 1) + \varepsilon u_2(\log u_2 - 1).$$

This regularization is motivated from the population model with diffusion matrix (14), see Section 2.2. The range of  $Dh_\varepsilon$  equals  $\mathbb{R}^2$ , as required in Hypothesis H1. However, this regularization makes necessary to regularize also the diffusion matrix,

$$A_\varepsilon(u) = A(u) + \varepsilon \begin{pmatrix} u_2 & 0 \\ 0 & u_1 \end{pmatrix}.$$

Then the regularized product  $D^2 h_\varepsilon(u) A_\varepsilon(u)$  is positive semi-definite (Hypothesis H2) only if  $s < 4$ . This restriction may be improved by developing a better regularization procedure.

This paper is organized as follows. In Section 2, we consider some examples of cross-diffusion systems, studied in the literature, and discuss the validity of our hypotheses. The general existence Theorem 2 is shown in Section 3. The proofs of Theorems 3 and 4 are presented in Sections 4 and 5, respectively. Some further results and open problems are discussed in Section 6. The appendix is concerned with some relations of our method to non-equilibrium thermodynamics; the derivation of a general population diffusion model, containing the diffusion matrices (10) and (13) as special cases; and the proof of a variant of the Aubin compactness lemma, needed in the proof of Theorem 3. To simplify the presentation, we write in the following  $L^p(0, T; H^k(\Omega))$  instead of  $L^p(0, T; H^k(\Omega; \mathbb{R}^n))$ .

## 2. EXAMPLES

We present some diffusion problems from physical and biological applications and discuss the validity of Hypotheses H1-H3. Some examples do not satisfy the hypotheses but similar ideas as detailed in the introduction apply.

### 2.1. Examples with volume filling.

*Volume-filling model of Burger.* Burger et al. [9] derived from a lattice-based hopping model a cross-diffusion system for  $n$  species, incorporating volume-filling effects and modeling the ion transport through narrow pores. The system consists of equation (1) with the diffusion matrix  $A(u) = (a_{ij}(u))$ , where

$$a_{ij}(u) = D_i u_i \quad \text{for } i \neq j, \quad a_{ii}(u) = D_i(1 - \rho + u_i), \quad \rho = \sum_{j=1}^n u_j,$$

and  $D_i > 0$  are some constants. It was shown in [9] that the entropy with density

$$h(u) = \sum_{i=1}^n u_i(\log u_i - 1) + (1 - \rho)(\log(1 - \rho) - 1),$$

where  $u = (u_1, \dots, u_n) \in D = \{u \in (0, \infty)^n : \sum_{i=1}^n u_i < 1\}$ , is a Lyapunov functional along solutions to (1) with  $f = 0$ . This entropy density satisfies Hypothesis H1 since  $u_i = ((Dh)^{-1}(w))_i = e^{w_i} / (1 + \sum_{j=1}^n e^{w_j}) \in D$  for  $w = (w_1, \dots, w_n) \in \mathbb{R}^n$ . The case of  $n = 2$  species was investigated in [8]. Then the diffusion matrix becomes

$$A(u) = \begin{pmatrix} D_1(1 - u_2) & D_1 u_1 \\ D_2 u_2 & D_2(1 - u_1) \end{pmatrix},$$

which is a special case of (10) with  $q(u_3) = u_3$  and  $\beta = D_2/D_1$ . A computation shows that for  $z = (z_1, z_2) \in \mathbb{R}^2$ ,

$$\begin{aligned} z^\top D^2 h(u) A(u) z &= D_1(1 - \rho) \left( \frac{z_1^2}{u_1} + \frac{z_2^2}{u_2} \right) + D_1 \frac{2 - \rho}{1 - \rho} (z_1 + z_2)^2 \\ &\quad + (D_2 - D_1) \frac{u_2}{1 - \rho} \left( z_1 + \frac{1 - u_1}{u_2} z_2 \right)^2. \end{aligned}$$

Consequently, assuming without loss of generality that  $D_2 > D_1$  (otherwise, change the indices 1 and 2), Hypothesis H2 is satisfied but not Hypothesis H2' since the quadratic form degenerates at  $\rho = 1$ . Burger et al. have shown in [8] that a global existence analysis is still possible. We generalize this result to general functions  $q$  in Theorem 3.

In the general case  $n > 2$  and the case of equal diffusivity constants  $D_1 = D_i$  for all  $i = 2, \dots, n$ , summing up all equations yields  $\partial_t \rho - \Delta \rho = 0$  in  $\Omega$  with homogeneous Neumann boundary conditions. Thus, by the classical maximum principle,  $\rho$  is bounded from below and above, and this yields  $L^\infty$  bounds for the components  $u_i \geq 0$ . The analysis of different diffusivity constants and  $n > 2$  is much more delicate and an open problem.

*Tumor-growth model.* Jackson and Byrne [32] derived a continuous mechanical model for the growth of symmetric avascular tumors in one space dimension. In this model, the mass balance equations for the volume fractions of the tumor cells  $u_1$ , the extracellular matrix  $u_2$ , and the water phases  $u_3 = 1 - u_1 - u_2$  are supplemented by equations for the velocity, obtained from a force balance. For a small cell-induced pressure coefficient (i.e.  $\theta = \beta = 1$  in [32]), the resulting diffusion matrix reads as

$$A(u) = \begin{pmatrix} u_1(1 - u_1) - u_1 u_2^2 & -u_1 u_2(1 + u_1) \\ -u_1 u_2 + u_2^2(1 - u_2) & u_2(1 - u_2)(1 + u_1) \end{pmatrix}.$$

With the entropy density (6), defined in  $D$  which is given by (7), we compute for  $z = (z_1, z_2)^\top \in \mathbb{R}^2$ ,

$$z^\top D^2 h(u) A(u) z = z_1^2 + (1 + u_1) z_2^2 + u_1 z_1 z_2 \geq \frac{1}{2} z_1^2 + \left(1 + u_1 - \frac{1}{2} u_2^2\right) z_2^2.$$

Thus, since  $u_2 \leq 1$ , Hypotheses H1, H2', and H2'' are satisfied and the global existence result follows from Theorem 3 if the reaction terms satisfy Hypothesis H3. This result was first proven in [33].

*Maxwell-Stefan equations.* The Maxwell-Stefan equations describe the diffusive transport of multicomponent gaseous mixtures. In the case of mixtures of ideal gases consisting of  $n$  components under isobaric, isothermal conditions with vanishing barycentric velocity, the equations read as

$$\partial_t u_i + \operatorname{div} J_i = f_i(u), \quad \nabla u_i = - \sum_{j \neq i} \frac{u_j J_i - u_i J_j}{D_{ij}}, \quad i = 1, \dots, n,$$

where  $D_{ij} > 0$  are the binary diffusion coefficients. The variables  $u_i$  are the molar concentrations of the mixture and they satisfy  $\sum_{j=1}^n u_j = 1$ . The inversion of the flux-gradient relations is not straightforward since the linear system in  $J_i$  has a singular matrix; see [4, 45]. The idea of [34] was to replace the last component  $u_n$  by the remaining ones by  $u_n = 1 - \sum_{j=1}^{n-1} u_j$  and to analyze the remaining  $n - 1$  equations. For  $n = 3$  components, the inversion can be made explicit, leading to system (1) for the remaining  $n - 1 = 2$  equations with the diffusion matrix

$$(17) \quad A(u) = \frac{1}{\delta(u)} \begin{pmatrix} d_2 + (d_0 - d_2)u_1 & (d_0 - d_1)u_1 \\ (d_0 - d_2)u_2 & d_1 + (d_0 - d_1)u_2 \end{pmatrix},$$

where we abbreviated  $d_{i+j-2} = D_{ij}$  and  $\delta(u) = d_1 d_2 (1 - u_1 - u_2) + d_0 (d_1 u_1 + d_2 u_2)$ . The diffusion matrix (5) in Section 1.2 is obtained after setting  $d_0 = 3$ ,  $d_1 = 2$ , and  $d_2 = 1$ .

Employing the entropy density (6) as in the previous example, we compute

$$z^\top D^2 h(u) A(u) z = \frac{d_2}{u_1} z_1^2 + \frac{d_1}{u_2} z_2^2 + \frac{d_0 u_1 u_2}{1 - u_1 - u_2} (z_1 + z_2)^2$$

for  $z = (z_1, z_2)^\top \in \mathbb{R}^2$ . Therefore, Hypothesis H1 and H2 are fulfilled with  $D$  given by (7). Moreover, Hypothesis H2' is satisfied with  $m_1 = m_2 = 0$ , and also Hypothesis H2'' holds. Theorem 2 yields the existence of global bounded weak solutions to this problem. It can be shown that this result holds true in the case of  $n > 3$  components; see [34].

## 2.2. Models without volume filling.

*Population model of Shigesada, Kawashima, and Teramoto.* The model describes the evolution of two population densities  $u_1$  and  $u_2$  governed by equation (1) with the diffusion matrix

$$A(u) = \begin{pmatrix} \alpha_{01} + 2\alpha_{11}u_1 + \alpha_{12}u_2 & \alpha_{12}u_1 \\ \alpha_{21}u_2 & \alpha_{20} + \alpha_{21}u_1 + 2\alpha_{22}u_2 \end{pmatrix},$$

where the coefficients  $\alpha_{ij}$  are nonnegative. Introducing the entropy density

$$h(u) = \frac{u_1}{\alpha_{12}} (\log u_1 - 1) + \frac{u_2}{\alpha_{21}} (\log u_2 - 1), \quad u = (u_1, u_2) \in D = (0, \infty)^2,$$

a computation shows that Hypothesis H2 is satisfied, i.e. for all  $z = (z_1, z_2)^\top \in \mathbb{R}^2$ ,

$$z^\top D^2 h(u) A(u) z = \left( \frac{\alpha_{11}}{\alpha_{12}} + \frac{\alpha_{10}}{\alpha_{12} u_1} \right) z_1^2 + \left( \frac{\alpha_{22}}{\alpha_{21}} + \frac{\alpha_{20}}{\alpha_{21} u_2} \right) z_2^2 + \left( \sqrt{\frac{u_2}{u_1}} z_1 + \sqrt{\frac{u_1}{u_2}} z_2 \right)^2.$$

Thus, we see that Hypothesis H2' with  $m_i = 0$  or  $m_i = 1$  and Hypothesis H2'' hold true. Since  $D = (0, \infty)^2$  is not bounded and the range of  $Dh$  is not the whole  $\mathbb{R}^2$ , we cannot apply Theorem 2. However, as the coefficients of  $A(u)$  depend only linearly on  $u$ , the proof of Theorem 4 can be adapted to this case and we conclude the existence of global nonnegative weak solutions to this problem. This statement was first proved in [14] using a different approximation procedure. In Theorem 4 we generalize this result.

*Semiconductor model with electron-hole scattering.* The carrier transport through a semiconductor device with strong electron-hole scattering but vanishing electric field can be modeled by equation (1) with diffusion matrix

$$A(u) = \frac{1}{1 + \mu_2 u_1 + \mu_1 u_2} \begin{pmatrix} \mu_1(1 + \mu_2 u_1) & \mu_1 \mu_2 u_1 \\ \mu_1 \mu_2 u_2 & \mu_2(1 + \mu_1 u_2) \end{pmatrix},$$

where  $u_1$  and  $u_2$  are the electron and hole densities, respectively, and  $\mu_1$  and  $\mu_2$  denote the corresponding (positive) mobility constants. This model was formally derived from the semiconductor Boltzmann equation with a collision operator taking into account strong electron-hole scattering [52]. The global existence of weak solutions was shown in [16].

Introducing the entropy density

$$h(u) = u_1(\log u_1 - 1) + u_2(\log u_2 - 1),$$

where  $u = (u_1, u_2) \in D = (0, \infty)^2$ , we obtain for  $z = (z_1, z_2)^\top \in \mathbb{R}^2$ ,

$$z^\top D^2 h(u) A(u) z = \frac{1}{1 + \mu_2 u_1 + \mu_1 u_2} \left( \frac{\mu_1}{u_1} z_1^2 + \frac{\mu_2}{u_2} z_2^2 + \mu_1 \mu_2 (z_1 + z_2)^2 \right).$$

Thus, Hypothesis H2 is satisfied but not H2' since the quadratic form is not uniformly positive. However, as shown in [16], bounds on the electron and hole masses (i.e. the integrals of  $u_1$  and  $u_2$ ) together with estimates from the entropy dissipation yield an  $H^1$  bound for  $u_1^{1/2}$  and  $u_2^{1/2}$ . Then, with the entropy variables  $w_i = \partial h / \partial u_i = \log u_i$ , the existence of global weak solutions was proved in [16].

### 3. PROOF OF THEOREM 2

The proof is based on the solution of a time-discrete and regularized problem, for which only Hypotheses H1, H2, and H3 are needed.

*Step 1. Solution of an approximate problem.* Let  $T > 0$ ,  $N \in \mathbb{N}$ ,  $\tau = T/N$ ,  $m \in \mathbb{N}$  with  $m > d/2$ , and  $\tau > 0$ . Let  $w^{k-1} \in L^\infty(\Omega; \mathbb{R}^n)$  be given. (If  $k = 1$ , we define  $w^0 = (Dh)^{-1}(u^0)$  which is possible since we assumed that  $u^0(x) \in D$  for  $x \in \Omega$ .) We wish to find  $w^k \in H^m(\Omega; \mathbb{R}^n)$  such that

$$(18) \quad \begin{aligned} & \frac{1}{\tau} \int_{\Omega} (u(w^k) - u(w^{k-1})) \cdot \phi dx + \int_{\Omega} \nabla \phi : B(w^k) \nabla w^k dx \\ & + \varepsilon \int_{\Omega} \left( \sum_{|\alpha|=m} D^\alpha w^k \cdot D^\alpha \phi + w^k \cdot \phi \right) dx = \int_{\Omega} f(u(w^k)) \cdot \phi dx \end{aligned}$$

for all  $\phi \in H^m(\Omega; \mathbb{R}^n)$ , where  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$  with  $|\alpha| = \alpha_1 + \dots + \alpha_n = m$  is a multiindex and  $D^\alpha = \partial^m / (\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n})$  is a partial derivative of order  $m$ , and  $u(w) = (Dh)^{-1}(w)$ . We observe that  $m$  is chosen in such a way that  $H^m(\Omega) \hookrightarrow L^\infty(\Omega)$ .

**Lemma 5** (Existence for the regularized system (18)). *Let Assumptions H1, H2, and H3 hold. Let  $u^0 \in L^\infty(\Omega; \mathbb{R}^n)$  be such that  $u^0(x) \in D$  for  $x \in \Omega$ , and let  $0 < \tau < 1/C_f$ , where  $C_f > 0$  is defined in H3. Then there exists a weak solution  $w^k \in H^m(\Omega; \mathbb{R}^n)$  to (18) satisfying  $u(w^k(x)) \in D$  for  $x \in \Omega$  and the discrete entropy-dissipation inequality*

$$(19) \quad \begin{aligned} & (1 - C_f \tau) \int_{\Omega} h(u(w^k)) dx + \tau \int_{\Omega} \nabla w^k : B(w^k) \nabla w^k dx + \varepsilon \tau \int_{\Omega} \left( \sum_{|\alpha|=m} |D^\alpha w|^2 + |w|^2 \right) dx \\ & \leq \tau C_f \text{meas}(\Omega) + \int_{\Omega} h(u(w^{k-1})) dx. \end{aligned}$$

*Proof.* Let  $y \in L^\infty(\Omega; \mathbb{R}^n)$  and  $\delta \in [0, 1]$  be given. We solve first the following linear problem:

$$(20) \quad a(w, \phi) = F(\phi) \quad \text{for all } \phi \in H^m(\Omega; \mathbb{R}^n),$$

where

$$\begin{aligned} a(w, \phi) &= \int_{\Omega} \nabla \phi : B(y) \nabla w dx + \varepsilon \int_{\Omega} \left( \sum_{|\alpha|=m} D^\alpha w \cdot D^\alpha \phi + w \cdot \phi \right) dx, \\ F(\phi) &= -\frac{\delta}{\tau} \int_{\Omega} (u(y) - u(w^{k-1})) \cdot \phi dx + \delta \int_{\Omega} f(u(y)) \cdot \phi dx. \end{aligned}$$

The forms  $a$  and  $F$  are bounded on  $H^m(\Omega; \mathbb{R}^n)$ . The matrix  $B(y) = A(u(y))(D^2 h)^{-1}(u(y))$  is positive semi-definite since, by Assumption H2,

$$z^\top B(y) z = ((D^2 h)^{-1} z)^\top (D^2 h) A(u(y)) ((D^2 h)^{-1} z) \geq 0 \quad \text{for all } z \in \mathbb{R}^n.$$

Hence, the bilinear form  $a$  is coercive:

$$a(w, w) \geq \varepsilon \int_{\Omega} \left( \sum_{|\alpha|=m} |D^\alpha w|^2 + |w|^2 \right) dx \geq \varepsilon C \|w\|_{H^m(\Omega)}^2 \quad \text{for } w \in H^m(\Omega; \mathbb{R}^n).$$

The last inequality follows from the generalized Poincaré inequality for  $H^m$  spaces [58, Chap. II.1.4, Formula (1.39)], and  $C > 0$  is some constant only depending on  $\Omega$ . Therefore, we can apply the Lax-Milgram lemma to obtain the existence of a unique solution  $w \in H^m(\Omega; \mathbb{R}^n) \hookrightarrow L^\infty(\Omega; \mathbb{R}^n)$  to (20). This defines the fixed-point operator  $S : L^\infty(\Omega; \mathbb{R}^n) \times [0, 1] \rightarrow L^\infty(\Omega; \mathbb{R}^n)$ ,  $S(y, \delta) = w$ , where  $w$  solves (20).

It holds  $S(y, 0) = 0$  for all  $y \in L^\infty(\Omega; \mathbb{R}^n)$ . We claim that the operator  $S$  is continuous. Indeed, let  $\delta_\eta \rightarrow \delta$  in  $\mathbb{R}$  and  $y_\eta \rightarrow y$  strongly in  $L^\infty(\Omega; \mathbb{R}^n)$  as  $\eta \rightarrow 0$  and set  $w_\eta = S(y_\eta, \delta_\eta) \in H^m(\Omega; \mathbb{R}^n)$ . Then, by continuity,  $u(y_\eta) \rightarrow u(y)$ ,  $B(y_\eta) \rightarrow B(y)$ , and  $f(u(y_\eta)) \rightarrow f(u(y))$  strongly in  $L^\infty(\Omega; \mathbb{R}^n)$ , and the above coercivity estimate gives a uniform bound for  $(w_\eta)$  in  $H^m(\Omega; \mathbb{R}^n)$ . This implies that, for a subsequence which is not relabeled,  $w_\eta \rightharpoonup w$  weakly in  $H^m(\Omega; \mathbb{R}^n)$ . Then, performing the limit  $\eta \rightarrow 0$  in (20), it follows that  $w = S(y, \delta)$ . In view of the compact embedding  $H^m(\Omega; \mathbb{R}^n) \hookrightarrow L^\infty(\Omega; \mathbb{R}^n)$ , we infer that for a subsequence,  $S(y_\eta, \delta_\eta) = w_\eta \rightarrow w = S(y, \delta)$  strongly in  $L^\infty(\Omega; \mathbb{R}^n)$ . Since the limit  $w$  is unique, the whole sequence  $(w_\eta)$  is converging which shows the continuity. Furthermore, the compact embedding  $H^m(\Omega; \mathbb{R}^n) \hookrightarrow L^\infty(\Omega; \mathbb{R}^n)$  yields the compactness of  $S$ .

It remains to prove a uniform bound for all fixed points of  $S(\cdot, \delta)$  in  $L^\infty(\Omega; \mathbb{R}^n)$ . Let  $w \in L^\infty(\Omega; \mathbb{R}^n)$  be such a fixed point. Then  $w$  solves (20) with  $y$  replaced by  $w$ . With the test function  $\phi = w$ , we find that

$$(21) \quad \begin{aligned} & \frac{\delta}{\tau} \int_{\Omega} (u(w) - u(w^{k-1})) \cdot w dx + \int_{\Omega} \nabla w : B(w) \nabla w dx \\ & + \varepsilon \int_{\Omega} \left( \sum_{|\alpha|=m} |D^\alpha w|^2 + |w|^2 \right) dx = \delta \int_{\Omega} f(u(w)) \cdot w dx. \end{aligned}$$

The convexity of  $h$  implies that  $h(x) - h(y) \leq Dh(x) \cdot (x - y)$  for all  $x, y \in D$ . Choosing  $x = u(w)$  and  $y = u(w^{k-1})$  and using  $Dh(u(w)) = w$ , this gives

$$\frac{\delta}{\tau} \int_{\Omega} (u(w) - u(w^{k-1})) \cdot w dx \geq \frac{\delta}{\tau} \int_{\Omega} (h(u(w)) - h(u(w^{k-1}))) dx.$$

Taking into account the positive semi-definiteness of  $B(w)$  and Assumption H3, (21) can be estimated as follows:

$$\begin{aligned} & \delta \int_{\Omega} h(u(w)) dx + \varepsilon \tau \int_{\Omega} \left( \sum_{|\alpha|=m} |D^\alpha w|^2 + |w|^2 \right) dx \\ & \leq C_f \tau \delta \int_{\Omega} (1 + h(u(w))) dx + \delta \int_{\Omega} h(u(w^{k-1})) dx. \end{aligned}$$

Choosing  $\tau < 1/C_f$ , this yields an  $H^m$  bound for  $w$  uniform in  $\delta$  (not uniform in  $\tau$  or  $\varepsilon$ ). The Leray-Schauder fixed-point theorem shows that there exists a solution  $w \in H^m(\Omega; \mathbb{R}^n)$  to (20) with  $y$  replaced by  $w$  and  $\delta = 1$ .  $\square$

*Step 2. Uniform bounds.* By Lemma 5, there exists a weak solution  $w^k \in H^m(\Omega; \mathbb{R}^n)$  to (18). Because of the boundedness of  $D$ ,  $a < u_i(w^k) < b$  for  $i = 1, \dots, n$ . We need a priori estimates uniform in  $\tau$  and  $\varepsilon$ . For this, let  $w^{(\tau)}(x, t) = w^k(x)$  and  $u^{(\tau)}(x, t) = u(w^k(x))$  for  $x \in \Omega$  and  $t \in ((k-1)\tau, k\tau]$ ,  $k = 1, \dots, N$ . At time  $t = 0$ , we set  $w^{(\tau)}(\cdot, 0) = Dh(u^0)$  and  $u^{(\tau)}(\cdot, 0) = u^0$ . Let  $u^{(\tau)} = (u_1^{(\tau)}, \dots, u_n^{(\tau)})$ . Furthermore, we introduce the shift operator  $(\sigma_\tau u^{(\tau)})(\cdot, t) = u(w^{k-1}(x))$  for  $x \in \Omega$ ,  $(k-1)\tau < t \leq k\tau$ ,  $k = 1, \dots, N$ . Then  $u^{(\tau)}$  solves the equation

$$(22) \quad \begin{aligned} & \frac{1}{\tau} \int_0^T \int_{\Omega} (u^{(\tau)} - \sigma_\tau u^{(\tau)}) \cdot \phi dx dt + \int_0^T \int_{\Omega} \nabla \phi : B(w^{(\tau)}) \nabla w^{(\tau)} dx dt \\ & + \varepsilon \int_0^T \int_{\Omega} \left( \sum_{|\alpha|=m} D^\alpha w^{(\tau)} \cdot D^\alpha \phi + w^{(\tau)} \cdot \phi \right) dx dt = \int_0^T \int_{\Omega} f(u^{(\tau)}) \cdot \phi dx dt \end{aligned}$$

for piecewise constant functions  $\phi : (0, T) \rightarrow H^m(\Omega; \mathbb{R}^n)$ . We note that this set of functions is dense in  $L^2(0, T; H^m(\Omega; \mathbb{R}^n))$  [53, Prop. 1.36], such that the weak formulation also holds for such functions. By Assumption H2' and  $\nabla w^{(\tau)} = D^2 h(u^{(\tau)}) \nabla u^{(\tau)}$ , we find that

$$\begin{aligned} \int_{\Omega} \nabla w^{(\tau)} : B(w^{(\tau)}) \nabla w^{(\tau)} dx &= \int_{\Omega} \nabla u^{(\tau)} : D^2 h(u^{(\tau)}) A(u^{(\tau)}) \nabla u^{(\tau)} dx \\ &\geq \sum_{i=1}^n \int_{\Omega} \alpha_i (u_i^{(\tau)})^2 |\nabla u_i^{(\tau)}|^2 dx = \sum_{i=1}^n \int_{\Omega} |\nabla \tilde{\alpha}_i(u_i^{(\tau)})|^2 dx, \end{aligned}$$

where  $\tilde{\alpha}'_i = \alpha_i$ . Then either  $\tilde{\alpha}_i(y) = (\alpha_i^*/m_i)(y-a)^{m_i}$  or  $\tilde{\alpha}_i(y) = (\alpha_i^*/m_i)(b-y)^{m_i}$ . It follows from (19) that

$$\begin{aligned} & (1-C_f\tau) \int_{\Omega} h(u(w^k))dx + \tau \sum_{i=1}^n \int_{\Omega} |\nabla \tilde{\alpha}_i(u_i(w^k))|^2 dx \\ & + \varepsilon\tau \int_{\Omega} \left( \sum_{|\alpha|=m} |D^\alpha w^k|^2 + |w^k|^2 \right) dx \leq C_f\tau \text{meas}(\Omega) + \int_{\Omega} h(u(w^{k-1}))dx. \end{aligned}$$

Solving these inequalities recursively, we infer that

$$\begin{aligned} & \int_{\Omega} h(u(w^k))dx + \tau \sum_{j=1}^k (1-C_f\tau)^{-(k+1-j)} \sum_{i=1}^n \int_{\Omega} |\nabla \tilde{\alpha}_i(u_i(w^j))|^2 dx \\ & + \varepsilon\tau \sum_{j=1}^k (1-C_f\tau)^{-(k+1-j)} \int_{\Omega} \left( \sum_{|\alpha|=m} |D^\alpha w^j|^2 + |w^j|^2 \right) dx \\ & \leq C_f\tau \text{meas}(\Omega) \sum_{j=1}^k (1-C_f\tau)^{-j} + \int_{\Omega} h(u^0)dx. \end{aligned}$$

Since  $(1-C_f\tau)^{-(k+1-j)} \geq 1$  for  $j = 1, \dots, k$  and  $\sum_{j=1}^k (1-C_f\tau)^{-j} \leq (\exp(C_fk\tau) + 1)/(C_f\tau)$ , we obtain for  $k\tau \leq T$ ,

$$\begin{aligned} & \int_{\Omega} h(u(w^k))dx + \tau \sum_{j=1}^k \sum_{i=1}^n \int_{\Omega} |\nabla \tilde{\alpha}_i(u_i(w^j))|^2 dx + \varepsilon\tau \sum_{j=1}^k \int_{\Omega} \left( \sum_{|\alpha|=m} |D^\alpha w^j|^2 + |w^j|^2 \right) dx \\ & \leq \text{meas}(\Omega)(1 + e^{C_fT}) + \int_{\Omega} h(u^0)dx. \end{aligned}$$

Together with the  $L^\infty$  bounds for  $u^{(\tau)}$ , this gives the following uniform bounds:

$$(23) \quad \|\tilde{\alpha}(u^{(\tau)})\|_{L^2(0,T;H^1(\Omega))} \leq C, \quad \sqrt{\varepsilon}\|w^{(\tau)}\|_{L^2(0,T;H^m(\Omega))} \leq C,$$

where  $C > 0$  denotes here and in the following a generic constant independent of  $\tau$  and  $\varepsilon$ . For  $m_i \leq 1$  (appearing in  $\tilde{\alpha}_i$ ), we have

$$\begin{aligned} & \|\nabla u_i^{(\tau)}\|_{L^2(0,T;L^2(\Omega))} = \|\alpha_i(u_i^{(\tau)})^{-1} \nabla \tilde{\alpha}_i(u_i^{(\tau)})\|_{L^2(0,T;L^2(\Omega))} \\ (24) \quad & \leq (\alpha_i^*)^{-1} (b-a)^{1-m_i} \|\nabla \tilde{\alpha}_i(u_i^{(\tau)})\|_{L^2(0,T;L^2(\Omega))} \leq C. \end{aligned}$$

If  $m_i > 1$ , it follows that either

$$\begin{aligned} (25) \quad & \|\nabla (u_i^{(\tau)} - a)^{m_i}\|_{L^2(0,T;L^2(\Omega))} = (\alpha_i^*)^{-1} \|\nabla \tilde{\alpha}_i(u_i^{(\tau)})\|_{L^2(0,T;L^2(\Omega))} \leq C \quad \text{or} \\ & \|\nabla (b - u_i^{(\tau)})^{m_i}\|_{L^2(0,T;L^2(\Omega))} = (\alpha_i^*)^{-1} \|\nabla \tilde{\alpha}_i(u_i^{(\tau)})\|_{L^2(0,T;L^2(\Omega))} \leq C. \end{aligned}$$

In order to derive a uniform estimate for the discrete time derivative, let  $\phi \in L^2(0,T;H^m(\Omega;\mathbb{R}^n))$ . Then

$$\begin{aligned} & \left| \frac{1}{\tau} \int_{\tau}^T \int_{\Omega} (u^{(\tau)} - \sigma_\tau u^{(\tau)}) \cdot \phi dx dt \right| \\ & \leq \|A(u^{(\tau)}) \nabla u^{(\tau)}\|_{L^2(0,T;L^2(\Omega))} \|\nabla \phi\|_{L^2(0,T;L^2(\Omega))} \\ & \quad + \varepsilon \|w^{(\tau)}\|_{L^2(0,T;H^m(\Omega))} \|\phi\|_{L^2(0,T;H^m(\Omega))} + \|f(u^{(\tau)})\|_{L^2(0,T;L^2(\Omega))} \|\phi\|_{L^2(0,T;L^2(\Omega))}. \end{aligned}$$

The first term on the right-hand side is uniformly bounded since, by Assumption H2",

$$\begin{aligned} \|(A(u^{(\tau)})\nabla u^{(\tau)})_i\|_{L^2(0,T;L^2(\Omega))}^2 &\leq \sum_{j=1, m_j > 1}^n \left\| \frac{a_{ij}(u^{(\tau)})}{\alpha_j(u_j^{(\tau)})} \right\|_{L^\infty(0,T;L^\infty(\Omega))}^2 \|\nabla \tilde{\alpha}_j(u_j^{(\tau)})\|_{L^2(0,T;L^2(\Omega))}^2 \\ &\quad + \sum_{j=1, m_j \leq 1}^n \|a_{ij}(u^{(\tau)})\|_{L^\infty(0,T;L^\infty(\Omega))}^2 \|\nabla u_j^{(\tau)}\|_{L^2(0,T;L^2(\Omega))}^2 \\ &\leq C, \end{aligned}$$

using (23) and (24). Thus, by the  $L^\infty$  bound for  $u^{(\tau)}$  and (23),

$$(26) \quad \frac{1}{\tau} \left| \int_\tau^T \int_\Omega (u^{(\tau)} - \sigma_\tau u^{(\tau)}) \cdot \phi dx dt \right| \leq C(\sqrt{\varepsilon} \|\phi\|_{L^2(0,T;H^m(\Omega))} + \|\phi\|_{L^2(0,T;H^1(\Omega))}),$$

which shows that

$$(27) \quad \tau^{-1} \|u^{(\tau)} - \sigma_\tau u^{(\tau)}\|_{L^2(\tau,T;(H^m(\Omega)))'} \leq C.$$

*Step 3. The limit  $(\tau, \varepsilon) \rightarrow 0$ .* The uniform estimates (27) and either (24) or (25) allow us to apply the Aubin lemma in the version of [25, Theorem 1] (if  $m_i \leq 1$ ) or in the version of [18, Theorem 3a] (if  $m_i > 1$ ; apply the theorem to  $v_i^{(\tau)} = u_i^{(\tau)} - a$  or  $v_i^{(\tau)} = b - u_i^{(\tau)}$ ), showing that, up to a subsequence which is not relabeled, as  $(\tau, \varepsilon) \rightarrow 0$ ,

$$u^{(\tau)} \rightarrow u \quad \text{strongly in } L^2(0,T;L^2(\Omega)) \text{ and a.e. in } \Omega \times (0,T).$$

Because of the boundedness of  $u_i^{(\tau)}$  in  $L^\infty$ , this convergence even holds in  $L^p(0,T;L^p(\Omega))$  for any  $p < \infty$ , which is a consequence of the dominated convergence theorem. In particular,  $A(u^{(\tau)}) \rightarrow A(u)$  strongly in  $L^p(0,T;L^p(\Omega))$ . Furthermore, by weak compactness, for a subsequence,

$$\begin{aligned} A(u^{(\tau)})\nabla u^{(\tau)} &\rightharpoonup U \quad \text{weakly in } L^2(0,T;L^2(\Omega)), \\ \tau^{-1}(u^{(\tau)} - \sigma_\tau u^{(\tau)}) &\rightharpoonup \partial_t u \quad \text{weakly in } L^2(0,T;H^m(\Omega)'), \\ \varepsilon w^{(\tau)} &\rightarrow 0 \quad \text{strongly in } L^2(0,T;H^m(\Omega)). \end{aligned}$$

We claim that  $U = A(u)\nabla u$ . Indeed, observing that  $\tilde{\alpha}_i(u_i^{(\tau)}) \rightarrow \tilde{\alpha}_i(u_i)$  a.e., it follows that

$$\nabla \tilde{\alpha}_i(u_i^{(\tau)}) \rightharpoonup \nabla \tilde{\alpha}_i(u_i) \quad \text{weakly in } L^2(0,T;L^2(\Omega)).$$

Let  $m_j > 1$ . Since the quotient  $a_{ij}(u^{(\tau)})/\alpha_j(u_j^{(\tau)})$  is bounded, by Assumption H2", a subsequence converges weakly\* in  $L^\infty$ . Taking into account the a.e. convergence of  $(u^{(\tau)})$ , this implies that  $a_{ij}(u^{(\tau)})/\alpha_j(u_j^{(\tau)}) \rightharpoonup^* a_{ij}(u)/\alpha_j(u_j)$  weakly\* in  $L^\infty(0,T;L^\infty(\Omega))$  and a.e. and hence strongly in any  $L^p(0,T;L^p(\Omega))$  for  $p < \infty$ . Furthermore, if  $m_j \leq 1$ , since  $a_{ij}$  is continuous in  $\overline{D}$  and  $u^{(\tau)}(x,t) \in D$ ,  $a_{ij}(u^{(\tau)}) \rightarrow a_{ij}(u)$  strongly in  $L^p(0,T;L^p(\Omega))$  for all  $p < \infty$ . We conclude that

$$\begin{aligned} (A(u^{(\tau)})\nabla u^{(\tau)})_i &= \sum_{j=1, m_j > 1}^n \frac{a_{ij}(u^{(\tau)})}{\alpha_j(u_j^{(\tau)})} \nabla \tilde{\alpha}_j(u_j^{(\tau)}) + \sum_{j=1, m_j \leq 1}^n a_{ij}(u^{(\tau)}) \nabla u_j^{(\tau)} \\ &\rightharpoonup \sum_{j=1, m_j \leq 1}^n \frac{a_{ij}(u)}{\alpha_j(u_j)} \nabla \tilde{\alpha}_j(u_j) + \sum_{j=1, m_j \leq 1}^n a_{ij}(u) \nabla u_j = (A(u)\nabla u)_i \end{aligned}$$



weakly in  $L^q(0, T; L^q(\Omega))$  for all  $q < 2$ . Thus proves the claim. Finally, we observe that  $f(u^{(\tau)}) \rightarrow f(u)$  strongly in  $L^p(0, T; L^p(\Omega))$  for  $p < \infty$ . Therefore, we can pass to the limit  $(\tau, \varepsilon) \rightarrow 0$  in (22) to obtain a solution to

$$\int_0^T \langle \partial_t u, \phi \rangle dt + \int_0^T \int_{\Omega} \nabla \phi : A(u) \nabla u dx dt = \int_0^T \int_{\Omega} f(u) \cdot \phi dx dt$$

for all  $\phi \in L^2(0, T; H^m(\Omega; \mathbb{R}^n))$ . In fact, performing the limit  $\varepsilon \rightarrow 0$  and then  $\tau \rightarrow 0$ , we see from (26) that  $\partial_t u \in L^2(0, T; H^1(\Omega)')$ , and consequently, the weak formulation also holds for all  $\phi \in L^2(0, T; H^1(\Omega))$ . Moreover, it contains the homogeneous Neumann boundary conditions in (2)

It remains to show that  $u(0)$  satisfies the initial datum. Let  $\tilde{u}^{(\tau)}$  be the linear interpolant  $\tilde{u}^{(\tau)}(t) = u_k - (k\tau - t)(u^k - u^{k-1})/\tau$  for  $(k-1)\tau \leq t \leq k\tau$ , where  $u^k = u(w^k)$ . Because of (27),

$$\|\partial_t \tilde{u}^{(\tau)}\|_{L^2(0, T-\tau; H^m(\Omega)')} \leq \tau^{-1} \|u^{(\tau)} - \sigma_{\tau} u^{(\tau)}\|_{L^2(\tau, T; H^m(\Omega)')} \leq C.$$

This shows that  $(\tilde{u}^{(\tau)})$  is bounded in  $H^1(0, T; H^m(\Omega)')$ . Thus, for a subsequence,  $\tilde{u}^{(\tau)} \rightharpoonup w$  weakly in  $H^1(0, T; H^m(\Omega)') \hookrightarrow C^0([0, T]; H^m(\Omega)')$  and, by weak continuity,  $\tilde{u}^{(\tau)}(0) \rightharpoonup w(0)$  weakly in  $H^m(\Omega; \mathbb{R}^n)'$ . However,  $\tilde{u}^{(\tau)}$  and  $u^{(\tau)}$  converge to the same limit since

$$\|\tilde{u}^{(\tau)} - u^{(\tau)}\|_{L^2(0, T-\tau; H^m(\Omega)')} \leq \|u^{(\tau)} - \sigma_{\tau} u^{(\tau)}\|_{L^2(\tau, T; H^m(\Omega)')} \leq \tau C \rightarrow 0$$

as  $\tau \rightarrow 0$ . We infer that  $w = u$  and  $u^0 = \tilde{u}^{(\tau)}(0) \rightharpoonup u(0)$  weakly in  $H^m(\Omega; \mathbb{R}^n)'$ . This shows that the initial datum is satisfied in the sense of  $H^m(\Omega; \mathbb{R}^n)'$  and, in view of  $u \in H^1(0, T; H^1(\Omega)')$ , also in  $H^1(\Omega; \mathbb{R}^n)'$ .

#### 4. PROOF OF THEOREM 3

We impose more general assumptions on  $q$  than in Theorem 3: Let  $q \in C^1([0, 1])$  be positive and nondecreasing on  $(0, 1)$  such that

$$(28) \quad q(0) = 0, \text{ there exists } \kappa > 0 \text{ such that } \frac{1}{2} y q'(y) \leq q(y) \leq \kappa q'(y)^2 \text{ for } y \in [0, 1].$$

For instance, the functions  $q(y) = y/(1+y)^s$  with  $0 \leq s \leq 1$  and  $q(y) = y^s$  with  $1 \leq s \leq 2$  fulfill (28). The case  $q(y) = y^s$  with any  $s \geq 1$  is also possible, see Theorem 9.

**Theorem 6.** *Let  $\beta > 0$  and let  $q \in C^1([0, 1])$  be positive and nondecreasing on  $(0, 1)$ . We assume that (28) holds and that  $A(u)$  is given by (10). Let  $u^0 = (u_1^0, u_2^0) \in L^1(\Omega)$  with  $0 < u_1^0, u_2^0 < 1$  and  $u_1^0 + u_2^0 < 1$ . Then there exists a solution  $u$  to (1)-(2) with  $f = 0$  satisfying*

$$\begin{aligned} 0 \leq u_i, u_3 \leq 1 \text{ in } \Omega, \quad t > 0, \quad \partial_t u_i \in L^2(0, T; H^1(\Omega)'), \\ q(u_3)^{1/2}, q(u_3)^{1/2} u_i \in L^2(0, T; H^1(\Omega)), \quad i = 1, 2, \end{aligned}$$

where  $u_3 = 1 - u_1 - u_2 \geq 0$ , the weak formulation (11) holds for all  $\phi \in L^2(0, T; H^1(\Omega))$ , and  $u(0) = u^0$  in the sense of  $H^1(\Omega; \mathbb{R}^2)'$ .

*Proof.* First we observe that we may assume that  $\beta \geq 1$  since otherwise we rescale the equations by  $t \mapsto \beta t$ , which removes this factor in the second equation but yields the factor  $1/\beta > 1$  in the first equation.

*Step 1: Verification of Assumptions H1-H2.* We claim that Assumptions H1 and H2 are satisfied for the entropy density

$$h(u) = u_1(\log u_1 - 1) + u_2(\log u_2 - 1) + \int_{1/2}^{u_3} \log q(s) ds$$

for  $u \in D = \{(u_1, u_2) \in (0, 1)^2 : u_1 + u_2 < 1\}$ . Indeed, the Hessian  $D^2h$  is positive definite on  $D$ . Therefore,  $h$  is strictly convex. We claim that  $Dh : D \rightarrow \mathbb{R}^2$  is invertible. For this, let  $(w_1, w_2) \in \mathbb{R}^2$  and define  $g(y) = (e^{w_1} + e^{w_2})q(1-y)$  for  $0 < y < 1$ . Since  $q$  is nondecreasing,  $g$  is nonincreasing. Furthermore,  $g(0) > 0$  and  $g(1) = 0$  using  $q(0) = 0$ . By continuity, there exists a unique fixed point  $0 < y_0 < 1$  such that  $g(y_0) = y_0$ . Then we define  $u_i = e^{w_i}q(1-y_0) > 0$  ( $i = 1, 2$ ) which satisfies  $u_1 + u_2 = g(y_0) = y_0 < 1$ . Consequently,  $u = (u_1, u_2) \in D$ . We set  $u_3 = 1 - u_1 - u_2 = 1 - y_0$  which gives  $w_i = \log(u_i/q(u_3)) = (\partial h / \partial u_i)(u)$ . Hence, Assumption H1 is satisfied.

To verify Assumption H2, we show the following lemma.

**Lemma 7** (Positive definiteness of  $(D^2h)A$ ). *The matrix  $(D^2h)A$  is positive semi-definite. Moreover, if  $yq'(y) \leq 2q(y)$  holds for all  $0 \leq y \leq 1$  (see (28)), it holds*

$$(29) \quad z^\top (D^2h)Az \geq \frac{q(u_3)}{u_1} z_1^2 + \frac{q(u_3)}{u_2} z_2^2 + \frac{q'(u_3)^2}{q(u_3)} (z_1 + z_2)^2 \quad \text{for } z = (z_1, z_2)^\top \in \mathbb{R}^2.$$

*Proof.* Set  $(D^2h)A = M_0 + (\beta - 1)M_1$ , where

$$M_0 = \begin{pmatrix} q(u_3)/u_1 & 0 \\ 0 & q(u_3)/u_2 \end{pmatrix} + \frac{q'(u_3)}{q(u_3)} (2q(u_3) + (u_1 + u_2)q'(u_3)) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

$$M_1 = \begin{pmatrix} 0 & q'(u_3) \\ q'(u_3) & q(u_3)/u_2 + 2q'(u_3) \end{pmatrix} + \frac{q'(u_3)^2 u_2}{q(u_3)} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

A straightforward computation shows that

$$z^\top M_0 z = \frac{q(u_3)}{u_1} z_1^2 + \frac{q(u_3)}{u_2} z_2^2 + \frac{q'(u_3)}{q(u_3)} (2q(u_3) + (1 - u_3)q'(u_3)) (z_1 + z_2)^2,$$

$$z^\top M_1 z = u_2 q(u_3) \left( \frac{q'(u_3)}{q(u_3)} z_1 + \left( \frac{1}{u_2} + \frac{q'(u_3)}{q(u_3)} \right) z_2 \right)^2 \geq 0$$

for all  $z = (z_1, z_2)^\top \in \mathbb{R}^2$ . If  $yq'(y) \leq 2q(y)$  holds, we find that  $2q(u_3) + (1 - u_3)q'(u_3) \geq q'(u_3)$  from which we infer the result.  $\square$

*Step 2: A priori estimates.* By Lemma 5, there exists a sequence  $(w^k)$  of weak solutions to (18) satisfying the entropy-dissipation inequality (19) with  $C_f = 0$ . Taking into account the identity  $B(w^k)\nabla w^k = A(u^k)\nabla u^k$ , where  $u^k = u(w^k)$ ,  $w^k$  is a solution to

$$(30) \quad \frac{1}{\tau} \int_{\Omega} (u^k - u^{k-1}) \cdot \phi dx + \int_{\Omega} \nabla \phi : A(u^k) \nabla u^k dx + \varepsilon \int_{\Omega} \left( \sum_{|\alpha|=m} D^\alpha w^k \cdot D^\alpha \phi + w^k \cdot \phi \right) dx = 0$$

for all  $\phi \in H^m(\Omega; \mathbb{R}^2)$ . Furthermore, because of

$$\nabla w^k : B(w^k) \nabla w^k = \nabla u^k : (D^2h)(u^k) A(u^k) \nabla u^k$$

and (29), the discrete entropy inequality (19) can be written as

$$\int_{\Omega} h(u^k) dx + 4\tau \int_{\Omega} q(u_3^k) \sum_{i=1}^2 |\nabla(u_i^k)^{1/2}|^2 dx + 4\tau \int_{\Omega} |\nabla q(u_3^k)^{1/2}|^2 dx$$

$$(31) \quad + \varepsilon \tau \int_{\Omega} \left( \sum_{|\alpha|=m} |D^{\alpha} w^k|^2 + |w^k|^2 \right) dx \leq \int_{\Omega} h(u^{k-1}) dx,$$

where  $u^k = (u_1^k, u_2^k)$  and  $u_3^k = 1 - u_1^k - u_2^k$ . Resolving the recursion, we infer that

$$(32) \quad \begin{aligned} & \int_{\Omega} h(u^k) dx + 4\tau \sum_{j=1}^k \int_{\Omega} q(u_3^j) \sum_{i=1}^2 |\nabla(u_i^j)^{1/2}|^2 dx + 4\tau \sum_{j=1}^k \int_{\Omega} |\nabla q(u_3^j)^{1/2}|^2 dx \\ & + \varepsilon \tau \sum_{j=1}^k \int_{\Omega} \left( \sum_{|\alpha|=m} |D^{\alpha} w^j|^2 + |w^j|^2 \right) dx \leq \int_{\Omega} h(u^0) dx. \end{aligned}$$

Using the generalized Poincaré inequality [58, Chap. II.1.4], we deduce that

$$(33) \quad \varepsilon \tau \sum_{j=1}^k \|w_i^j\|_{H^m(\Omega)}^2 \leq C, \quad i = 1, 2.$$

The assumption  $q(y) \leq \kappa q'(y)^2$  (see (28)) implies an  $H^1$  estimate for  $u_3^k$ :

$$\tau \sum_{j=1}^k \int_{\Omega} |\nabla u_3^j|^2 dx \leq \kappa \tau \sum_{j=1}^k \int_{\Omega} \frac{q'(u_3^j)^2}{q(u_3^j)} |\nabla u_3^j|^2 dx dt = 4\kappa \tau \sum_{j=1}^k \int_{\Omega} |\nabla q(u_3^j)^{1/2}|^2 dx dt \leq C.$$

The boundedness of  $(u_i^k)$  in  $L^{\infty}$  yields an  $L^2$  estimate for  $(u_i^k)$  and hence, also for  $(u_3^k)$ . Therefore,

$$(34) \quad \tau \sum_{j=1}^k \|u_3^j\|_{H^1(\Omega)}^2 \leq C.$$

We cannot perform the simultaneous limit  $(\tau, \varepsilon) \rightarrow 0$  as in the proof of Theorem 2 since we need a compactness result of the type of Lemma 13 (see Appendix C) in which the discrete time derivative is uniformly bounded in  $H^1(\Omega)'$  and not in the larger space  $H^m(\Omega)'$  with  $m > d/2$ . Therefore, we pass to the limit in two steps.

*Step 3: Limit  $\varepsilon \rightarrow 0$ .* We fix  $k \in \{1, \dots, N\}$  and set  $u_i^{(\varepsilon)} = u_i^k$  for  $i = 1, 2, 3$ ,  $w_i^{(\varepsilon)} = w_i^k$  for  $i = 1, 2$ , and  $u^{(\varepsilon)} = (u_1^{(\varepsilon)}, u_2^{(\varepsilon)})$ . The uniform  $L^{\infty}$  bounds for  $(u_i^{(\varepsilon)})$  and estimates (33)-(34) as well as the compact embedding  $H^1(\Omega) \hookrightarrow L^2(\Omega)$  show that there exist subsequences which are not relabeled such that, as  $\varepsilon \rightarrow 0$ ,

$$\begin{aligned} u_i^{(\varepsilon)} &\rightharpoonup^* u_i \quad \text{weakly}^* \text{ in } L^{\infty}(\Omega), \quad i = 1, 2, \\ u_3^{(\varepsilon)} &\rightharpoonup u_3 \quad \text{weakly in } H^1(\Omega), \\ u_3^{(\varepsilon)} &\rightarrow u_3 \quad \text{strongly in } L^2(\Omega), \\ \varepsilon w_i^{(\varepsilon)} &\rightarrow 0 \quad \text{strongly in } H^m(\Omega). \end{aligned}$$

Moreover, since  $u_3^{(\varepsilon)} = 1 - u_1^{(\varepsilon)} - u_2^{(\varepsilon)}$ , we find that in the limit  $\varepsilon \rightarrow 0$ ,  $u_3 = 1 - u_1 - u_2$ . Since  $u_3^{(\varepsilon)} \rightarrow u_3$  a.e. in  $\Omega$  and  $q$  is continuous, we have  $q(u_3^{(\varepsilon)})^{1/2} \rightarrow q(u_3)^{1/2}$  a.e. in  $\Omega$ . The sequence  $(q(u_3^{(\varepsilon)})^{1/2})$  is bounded in  $L^{\infty}(\Omega)$ . Therefore, by dominated convergence,

$$(35) \quad q(u_3^{(\varepsilon)})^{1/2} \rightarrow q(u_3)^{1/2} \quad \text{strongly in } L^p(\Omega), \quad p < \infty.$$

Thus, the  $L^2$  bound for  $(\nabla q(u_3^{(\varepsilon)})^{1/2})$  from (32) shows that, up to a subsequence,

$$(36) \quad \nabla(q(u_3^{(\varepsilon)})^{1/2}) \rightharpoonup \nabla(q(u_3)^{1/2}) \quad \text{weakly in } L^2(\Omega).$$

Furthermore, in view of

$$\nabla(q(u_3^{(\varepsilon)})^{1/2} u_i^{(\varepsilon)}) = u_i^{(\varepsilon)} \nabla(q(u_3^{(\varepsilon)})^{1/2}) + 2(u_i^{(\varepsilon)})^{1/2} q(u_3^{(\varepsilon)})^{1/2} \nabla((u_i^{(\varepsilon)})^{1/2}),$$

estimate (32) and the  $L^\infty$  bounds show that for  $i = 1, 2$ ,

$$(37) \quad \begin{aligned} \int_{\Omega} |\nabla(q(u_3^{(\varepsilon)})^{1/2} u_i^{(\varepsilon)})|^2 dx &\leq 2 \int_{\Omega} (u_i^{(\varepsilon)})^2 |\nabla(q(u_3^{(\varepsilon)})^{1/2})|^2 dx + 8 \int_{\Omega} u_i^{(\varepsilon)} q(u_3^{(\varepsilon)}) |\nabla(u_i^{(\varepsilon)})^{1/2}|^2 dx \\ &\leq C \int_{\Omega} |\nabla(q(u_3^{(\varepsilon)})^{1/2})|^2 dx + C \int_{\Omega} q(u_3^{(\varepsilon)}) |\nabla(u_i^{(\varepsilon)})^{1/2}|^2 dx \leq C. \end{aligned}$$

It follows that

$$(38) \quad \|q(u_3^{(\varepsilon)})^{1/2} u_i^{(\varepsilon)}\|_{H^1(\Omega)} \leq C.$$

We conclude, again up to a subsequence, that  $q(u_3^{(\varepsilon)})^{1/2} u_i^{(\varepsilon)}$  converges weakly in  $H^1(\Omega)$  and, because of the compact embedding  $H^1(\Omega) \hookrightarrow L^2(\Omega)$ , strongly in  $L^2(\Omega)$  (and in fact, in every  $L^p(\Omega)$ ) to some function  $y \in H^1(\Omega)$ . As  $q(u_3^{(\varepsilon)})^{1/2} \rightarrow q(u_3)^{1/2}$  strongly in  $L^p(\Omega)$  and  $u_i^{(\varepsilon)} \rightharpoonup u_i$  weakly in  $L^p(\Omega)$  for all  $p < \infty$ , we obtain  $q(u_3^{(\varepsilon)})^{1/2} u_i^{(\varepsilon)} \rightharpoonup q(u_3)^{1/2} u_i$  weakly in  $L^p(\Omega)$ . Consequently,  $y = q(u_3)^{1/2} u_i$  and

$$(39) \quad q(u_3^{(\varepsilon)})^{1/2} u_i^{(\varepsilon)} \rightarrow q(u_3)^{1/2} u_i \quad \text{strongly in } L^p(\Omega), \quad p < \infty,$$

$$(40) \quad \nabla(q(u_3^{(\varepsilon)})^{1/2} u_i^{(\varepsilon)}) \rightharpoonup \nabla(q(u_3)^{1/2} u_i) \quad \text{weakly in } L^2(\Omega).$$

We wish to pass to the limit  $\varepsilon \rightarrow 0$  in

$$\begin{aligned} (A(u^{(\varepsilon)}) \nabla u^{(\varepsilon)})_i &= q(u_3^{(\varepsilon)}) \nabla u_i^{(\varepsilon)} + u_i^{(\varepsilon)} q'(u_3^{(\varepsilon)}) \nabla(1 - u_3^{(\varepsilon)}) \\ &= q(u_3^{(\varepsilon)})^{1/2} \nabla(q(u_3^{(\varepsilon)})^{1/2} u_i^{(\varepsilon)}) - 3q(u_3^{(\varepsilon)})^{1/2} u_i^{(\varepsilon)} \nabla(q(u_3^{(\varepsilon)})^{1/2}). \end{aligned}$$

Taking into account (35) and (40), the first summand converges weakly in  $L^1(\Omega)$  to  $q(u_3)^{1/2} \times \nabla(q(u_3)^{1/2} u_i)$ . By (36) and (39), the second summand converges weakly in  $L^1(\Omega)$  to  $-3q(u_3)^{1/2} u_i \nabla(q(u_3)^{1/2})$ . Thus, performing the limit  $\varepsilon \rightarrow 0$  in (30) and setting  $u_i^k := u_i$  ( $i = 1, 2, 3$ ), we find that  $u = (u_1, u_2)$  solves

$$\frac{1}{\tau} \int_{\Omega} (u^k - u^{k-1}) \cdot \phi dx + \sum_{i=1}^2 \beta_i \int_{\Omega} q(u_3^k)^{1/2} (\nabla(q(u_3^k)^{1/2} u_i^k) - 3u_i^k \nabla(q(u_3^k)^{1/2})) \cdot \nabla \phi_i dx = 0$$

for all  $\phi = (\phi_1, \phi_2) \in H^m(\Omega; \mathbb{R}^2)$ , where we recall that  $\beta_1 = 1$  and  $\beta_2 = \beta$ . By a density argument, this equation also holds for all  $\phi \in H^1(\Omega; \mathbb{R}^2)$ . Because of the weak convergence (40) and estimate (37), the limit  $q(u_3^k)^{1/2} u_i^k$  satisfies the bound

$$(41) \quad \tau \sum_{j=1}^k \|q(u_3^j)^{1/2} u_i^j\|_{H^1(\Omega)}^2 \leq C,$$

where  $C > 0$  depends on  $T_k = k\tau$  but not on  $\tau$ . Furthermore, by (32) and (34),

$$(42) \quad \tau \sum_{j=1}^k \|u_3^j\|_{H^1(\Omega)}^2 \leq C.$$

*Step 4: Limit  $\tau \rightarrow 0$ .* Let  $u^{(\tau)}(x, t) = u^k(x)$  for  $x \in \Omega$  and  $t \in ((k-1)\tau, k\tau]$ ,  $k = 1, \dots, N$ . At time  $t = 0$ , we set  $u^{(\tau)}(\cdot, 0) = u^0$ . Then  $u^{(\tau)}$  solves the equation

$$(43) \quad \begin{aligned} & \frac{1}{\tau} \int_{\tau}^T \int_{\Omega} (u^{(\tau)} - \sigma_{\tau} u^{(\tau)}) \phi dx dt \\ & + \sum_{i=1}^2 \beta_i \int_{\tau}^T \int_{\Omega} q(u_3^{(\tau)})^{1/2} (\nabla(q(u_3^{(\tau)})^{1/2} u_i^{(\tau)}) - 3u_i^{(\tau)} \nabla(q(u_3^{(\tau)})^{1/2})) \cdot \nabla \phi_i dx dt = 0 \end{aligned}$$

for all  $\phi = (\phi_1, \phi_2) \in L^2(0, T; H^1(\Omega; \mathbb{R}^2))$  and estimates (41) and (42) give

$$(44) \quad \|q(u_3^{(\tau)})^{1/2} u_i^{(\tau)}\|_{L^2(0, T; H^1(\Omega))} + \|u_3^{(\tau)}\|_{L^2(0, T; H^1(\Omega))} \leq C, \quad i = 1, 2,$$

where  $C > 0$  is independent of  $\tau$ . Clearly, we still have the  $L^\infty$  bounds:

$$\|u_i^{(\tau)}\|_{L^\infty(0, T; L^\infty(\Omega))} + \|u_3^{(\tau)}\|_{L^\infty(0, T; L^\infty(\Omega))} \leq C.$$

We claim that the discrete time derivative of  $u_3^{(\tau)}$  is also uniformly bounded. Indeed, since

$$\begin{aligned} & \int_0^T \|q(u_3^{(\tau)})^{1/2} [\nabla(q(u_3^{(\tau)})^{1/2} u_i^{(\tau)}) - 3u_i^{(\tau)} \nabla(q(u_3^{(\tau)})^{1/2})]\|_{L^2(\Omega)}^2 dt \\ & \leq C \int_0^T \|\nabla(q(u_3^{(\tau)})^{1/2} u_i^{(\tau)})\|_{L^2(\Omega)}^2 dt + C \int_0^T \|\nabla(q(u_3^{(\tau)})^{1/2})\|_{L^2(\Omega)}^2 dt \leq C, \end{aligned}$$

we find that

$$(45) \quad \begin{aligned} & \tau^{-1} \|u_i^{(\tau)} - \sigma_{\tau} u_i^{(\tau)}\|_{L^2(0, T; H^1(\Omega)')} \\ & \leq C \|q(u_3^{(\tau)})^{1/2} (\nabla(q(u_3^{(\tau)})^{1/2} u_i^{(\tau)}) - 3u_i^{(\tau)} \nabla(q(u_3^{(\tau)})^{1/2}))\|_{L^2(0, T; L^2(\Omega))} \leq C, \end{aligned}$$

which immediately yields

$$(46) \quad \tau^{-1} \|u_3^{(\tau)} - \sigma_{\tau} u_3^{(\tau)}\|_{L^2(\tau, T; H^1(\Omega)')} \leq C.$$

Then (44) and (46) allow us to apply the Aubin lemma in the version of [25] to conclude (up to a subsequence) the convergence

$$u_3^{(\tau)} \rightarrow u_3 \quad \text{strongly in } L^2(0, T; L^2(\Omega))$$

as  $\tau \rightarrow 0$ . Since  $(q(u_3^{(\tau)})^{1/2})$  is bounded in  $L^\infty$  and  $q(u_3^{(\tau)})^{1/2} \rightarrow q(u_3)^{1/2}$  a.e., the dominated convergence theorem implies that

$$q(u_3^{(\tau)})^{1/2} \rightarrow q(u_3)^{1/2} \quad \text{strongly in } L^p(0, T; L^p(\Omega)), \quad p < \infty.$$

In particular,  $(q(u_3^{(\tau)})^{1/2})$  is relatively compact in  $L^2(0, T; L^2(\Omega))$ . Furthermore, by (44) and (45),

$$\begin{aligned} & q(u_3^{(\tau)})^{1/2} \rightharpoonup q(u_3)^{1/2} \quad \text{weakly in } L^2(0, T; H^1(\Omega)), \\ & \tau^{-1} (u_i^{(\tau)} - \sigma_{\tau} u_i^{(\tau)}) \rightharpoonup \partial_t u_i \quad \text{weakly in } L^2(0, T; H^1(\Omega)'), \quad i = 1, 2. \end{aligned}$$

Since  $(u_i^{(\tau)})$  converges weakly in  $L^p(0, T; L^p(\Omega))$  and  $(q(u_3^{(\tau)})^{1/2})$  converges strongly in  $L^p(0, T; L^p(\Omega))$  for all  $p < \infty$ , we have

$$q(u_3^{(\tau)})^{1/2} u_i^{(\tau)} \rightharpoonup q(u_3)^{1/2} u_i \quad \text{weakly in } L^p(0, T; L^p(\Omega)), \quad p < \infty.$$

In fact, we can prove that this convergence is even strong. Indeed, applying Lemma 13 in Appendix C to  $y_\tau = q(u_3^{(\tau)})^{1/2}$  and  $z_\tau = u_i^{(\tau)}$ , we infer that, up to a subsequence,

$$q(u_3^{(\tau)})^{1/2} u_i^{(\tau)} \rightarrow q(u_3)^{1/2} u_i \quad \text{strongly in } L^p(0, T; L^p(\Omega)), \quad p < \infty.$$

The above convergence results show that

$$\begin{aligned} & q(u_3^{(\tau)})^{1/2} \nabla(q(u_3^{(\tau)})^{1/2} u_i^{(\tau)}) - 3q(u_3^{(\tau)})^{1/2} u_i^{(\tau)} \nabla(q(u_3^{(\tau)})^{1/2}) \\ & \rightharpoonup q(u_i)^{1/2} \nabla(q(u_3)^{1/3} u_i) - 3q(u_3)^{1/2} u_i \nabla(q(u_3)^{1/2}) \end{aligned}$$

weakly in  $L^1(0, T; L^1(\Omega))$ . Thus, passing to the limit  $\tau \rightarrow 0$  in (43) yields (11). This finishes the proof.  $\square$

**Remark 8.** We may allow for nonvanishing reaction terms  $f(u)$  in (1) if  $f$  depends linearly on  $u_1$  and  $u_2$  and (possibly) nonlinearly on  $u_3 = 1 - u_1 - u_2$ , since in the above proof  $(u_1^{(\tau)})$  and  $(u_2^{(\tau)})$  converge weakly in  $L^p$ , whereas  $(u_3^{(\tau)})$  converges strongly in  $L^p$  ( $p < \infty$ ).  $\square$

Functions  $q(y) = y^s$  with  $s > 2$  can be also considered.

**Theorem 9.** Let  $\beta > 1$  and  $q(y) = y^s$  with  $s \geq 1$ . Let  $u^0 \in L^1(\Omega)$  with  $0 < u^0 < 1$  in  $\Omega$ . Then there exists a weak solution to (1)-(2) satisfying the weak formulation (11),  $u(0) = u^0$  in  $H^1(\Omega; \mathbb{R}^2)'$ , and

$$\begin{aligned} & 0 \leq u_i, u_3 \leq 1 \text{ in } \Omega, \quad t > 0, \quad \partial_t u_i \in L^2(0, T; H^1(\Omega)'), \\ & u_3^{\alpha/2}, u_3^{\alpha/2} u_i \in L^2(0, T; H^1(\Omega)), \quad i = 1, 2. \end{aligned}$$

*Proof.* It is sufficient to consider the case  $s > 2$  since the case  $s \leq 2$  is contained in Theorem 6. Assumptions H1 and H2 hold since this requires only  $q(0) = 0$  and  $q'(s) \geq 0$ . By Lemma 5, there exists a sequence  $(w^k)$  of weak solutions to (18) satisfying the entropy-dissipation inequality (31), which reads as

$$\begin{aligned} & \int_{\Omega} h(u^k) dx + 4\tau \int_{\Omega} (u_3^k)^s \sum_{i=1}^2 |\nabla(u_i^k)^{1/2}|^2 dx + s\tau \int_{\Omega} ((2-s)u_3^k + s)(u_3^k)^{s-2} |\nabla u_3^k|^2 dx \\ & + \varepsilon\tau \int_{\Omega} \left( \sum_{|\alpha|=m} |D^\alpha w^k|^2 + |w^k|^2 \right) dx \leq \int_{\Omega} h(u^{k-1}) dx. \end{aligned}$$

Observing that

$$\begin{aligned} & ((2-s)u_3^k + s)(u_3^k)^{s-2} |\nabla u_3^k|^2 = \frac{4}{s^2} (2u_3^k + s(1-u_3^k)) |\nabla(u_3^k)^{s/2}|^2 \\ & \geq \frac{4}{s^2} \min\{2, s\} |\nabla(u_3^k)^{s/2}|^2, \end{aligned}$$

we infer after summation the uniform estimates

$$\tau \sum_{j=1}^k \sum_{i=1}^2 \|(u_3^j)^{s/2} \nabla(u_i^j)^{1/2}\|_{L^2(\Omega)}^2 + \tau \sum_{j=1}^k \|(u_3^j)^{s/2}\|_{H^1(\Omega)}^2 + \varepsilon\tau \sum_{j=1}^k \|w^j\|_{H^1(\Omega)}^2 \leq C.$$

Compared to the proof of Theorem 2, we cannot conclude an  $H^1$  bound for  $u_3^k$  but only for  $(u_3^k)^{s/2}$ . Set  $u_i^{(\varepsilon)} = u_i^k$  for  $i = 1, 2, 3$ . The function  $z \mapsto z^{2/s}$  for  $z \geq 0$  is Hölder continuous since  $s > 2$ . Therefore, by the lemma of Chavent and Jaffre [13, p. 141],

$$\|u_3^{(\varepsilon)}\|_{W^{2/s, s}(\Omega)} \leq C \|(u_3^{(\varepsilon)})^{s/2}\|_{H^1(\Omega)} \leq C.$$

Since the embedding  $W^{2/s,s}(\Omega) \hookrightarrow L^s(\Omega)$  is compact, we conclude the existence of a subsequence (not relabeled) such that

$$u_3^{(\varepsilon)} \rightarrow u_3 \quad \text{strongly in } L^s(\Omega).$$

At this point, we can proceed as in Step 3 of the proof of Theorem 2.

For the limit  $\tau \rightarrow 0$ , we need another compactness argument. Let  $u_i^{(\tau)}$  be defined as in the proof of Theorem 2. We have the uniform estimates

$$\begin{aligned} \|(u_3^{(\tau)})^{s/2} u_i^{(\tau)}\|_{L^2(0,T;H^1(\Omega))} + \|(u_3^{(\tau)})^{s/2}\|_{L^2(0,T;H^1(\Omega))} &\leq C, \\ \tau^{-1} \|u_3^{(\tau)} - \sigma_\tau u_3^{(\tau)}\|_{L^2(0,T;H^1(\Omega)')} &\leq C. \end{aligned}$$

Then, by the generalization of the Aubin lemma in [18, Theorem 3], we obtain, up to a subsequence, as  $\tau \rightarrow 0$ ,

$$u_3^{(\tau)} \rightarrow u_3 \quad \text{strongly in } L^s(0,T;L^s(\Omega)).$$

The remaining proof is exactly as Step 4 of the proof of Theorem 2.  $\square$

**Remark 10** (Generalization of Theorem 9). Let  $q \in C^1([0,1])$  be positive and nondecreasing on  $(0,1)$  such that there exist  $0 < \kappa_0 < 1$  and  $\kappa_1 > 0$  such that for all  $0 \leq z \leq 1$ ,

$$2q(y) \geq (y - 1 + \kappa_0)q'(y) \quad \text{and either } q(y) \leq \kappa_1 q'(y)^2 \text{ or } q(y) \leq \kappa_1 y^{2-s} q'(y)^2,$$

where  $s > 2$ . This includes  $q(y) = y^s$  with  $s > 2$  ( $0 \leq y \leq 1$ ). Indeed, let  $\kappa_0 = 2/s < 1$  and  $\kappa_1 = 1/s^2$ . Then  $2q(y) \geq (y - 1 + \kappa_0)q'(y)$  is equivalent to  $y \leq 1$  which is true, and  $q(y) = y^s \leq \kappa_1 y^{2-s} q'(y)^2 = y^s$ . The third term of the entropy-dissipation inequality (31) can be estimated as

$$\begin{aligned} \int_{\Omega} \frac{q'(u_3^k)}{q(u_3^k)} (2q(u_3^k) + (1 - u_3^k)q'(u_3^k)) |\nabla u_3^k|^2 dx &\geq \kappa_0 \int_{\Omega} \frac{q'(u_3^k)^2}{q(u_3^k)} |\nabla u_3^k|^2 dx \\ &= 4\kappa_0 \int_{\Omega} |\nabla q(u_3^k)^{1/2}|^2 dx, \end{aligned}$$

and we conclude the estimate

$$\tau \sum_{j=1}^k \|q(u_3^j)^{1/2}\|_{H^1(\Omega)} \leq C.$$

In case  $q(y) \leq \kappa_1 q'(y)^2$ , it follows that

$$\begin{aligned} \tau \sum_{j=1}^k \int_{\Omega} |\nabla u_3^j|^2 dx dt &\leq \kappa_1 \tau \sum_{j=1}^k \int_{\Omega} \frac{q'(u_3^j)^2}{q(u_3^j)} |\nabla u_3^j|^2 dx dt \\ &= 4\kappa_1 \tau \sum_{j=1}^k \int_{\Omega} |\nabla q(u_3^j)^{1/2}|^2 dx dt \leq C. \end{aligned}$$

In the other case  $q(y) \leq \kappa_1 y^{2-s} q'(y)^2$ , we find that

$$\tau \sum_{j=1}^k \int_{\Omega} |\nabla (u_3^j)^{s/2}|^2 dx dt \leq \frac{s^2}{4} \kappa_1 \tau \sum_{j=1}^k \int_{\Omega} \frac{q'(u_3^j)^2}{q(u_3^j)} |\nabla u_3^j|^2 dx dt \leq C.$$

This allows us to apply the generalized Aubin lemma and we proceed as in the above proof.  $\square$

## 5. PROOF OF THEOREM 4

We prove a slightly more general result than stated in Theorem 4 by allowing for more general functions  $p_1$  and  $p_2$ . We suppose that

$$(47) \quad p_i(u) = \alpha_{i0} + a_{i1}(u_1) + a_{i2}(u_2), \quad i = 1, 2,$$

where  $\alpha_{10}, \alpha_{20} > 0$  are positive numbers,  $a_{12}, a_{21}$  are continuously differentiable, and  $a_{11}, a_{22}$  are continuous functions on  $[0, \infty)$ . Assume that there exist constants  $\alpha_{ij} > 0$  such that

$$(48) \quad a_{11}(y) \geq \alpha_{11}y^s, \quad a_{22}(y) \geq \alpha_{22}y^s, \quad a'_{12}(y) \geq \alpha_{12}y^{s-1}, \quad a'_{21}(y) \geq \alpha_{21}y^{s-1}$$

for all  $y \geq 0$ , where  $1 < s < 4$ , and that there exist  $C > 0$  and  $\sigma < 2s(1 + 1/d) - 1$  such that for all  $u = (u_1, u_2) \in (0, \infty)^2$  and  $i = 1, 2$ , it holds that  $p_i(u) \leq C(1 + |u_1|^\sigma + |u_2|^\sigma)$ .

Since the entropy density  $h(u)$ , defined in (16), may not fulfill Hypothesis H1, we need to regularize:

$$h_\varepsilon(u) = h(u) + \varepsilon(u_1(\log u_1 - 1) + u_2(\log u_2 - 1)), \quad \varepsilon > 0.$$

Then

$$Dh_\varepsilon(u) = Dh(u) + \varepsilon(\log u_1, \log u_2)^\top,$$

and the range of  $Dh_\varepsilon$  is  $\mathbb{R}^2$ . The Hessian of  $h_\varepsilon$  equals

$$(49) \quad H_\varepsilon = D^2h_\varepsilon(u) = \begin{pmatrix} a'_{21}(u_1)/u_1 + \varepsilon/u_1 & 0 \\ 0 & a'_{12}(u_2)/u_2 + \varepsilon/u_2 \end{pmatrix},$$

showing that each component of  $Dh_\varepsilon$  is strictly increasing, and thus,  $(Dh_\varepsilon)^{-1} : \mathbb{R}^2 \rightarrow D = (0, \infty)^2$  is well defined. Hence,  $h_\varepsilon$  fulfills Hypothesis H1. We also regularize the diffusion matrix by setting

$$(50) \quad A_\varepsilon(u) = A(u) + \varepsilon \begin{pmatrix} u_2 & 0 \\ 0 & u_1 \end{pmatrix}.$$

The entropy density is regularized similarly in [23] but the diffusion matrix is regularized here in a different way.

*Step 1: Verification of Hypotheses H1 and H2.* Set  $H = D^2h(u)$ . First, we show that  $HA$  and  $H_\varepsilon A_\varepsilon$  are positive definite under additional conditions on the functions  $a_{ij}$ .

**Lemma 11** (Positive definiteness of  $HA$  and  $H_\varepsilon A_\varepsilon$ ). *Let  $H = D^2h(u)$ , where  $h$  is defined by (16), and let  $A = A(u)$  be given by (13) with  $p_i$  as in (47). If for some  $s \geq 1$ ,  $sa_{12}(u_2) \geq u_2a'_{12}(u_2)$ ,  $sa_{21}(u_1) \geq u_1a'_{21}(u_1)$ , and  $a'_{11}(u_1)a'_{22}(u_2) \geq (1 - 1/s)a'_{12}(u_2)a'_{21}(u_1)$  for all  $u_1, u_2 \geq 0$ , then  $HA$  is positive definite and for all  $z = (z_1, z_2)^\top \in \mathbb{R}^2$ ,*

$$z^\top H A z \geq \frac{\alpha_{10} + a_{11}(u_1)}{u_1} a'_{21}(u_1) z_1^2 + \frac{\alpha_{20} + a_{22}(u_2)}{u_2} a'_{12}(u_2) z_2^2.$$

*Furthermore, if additionally  $4a_{21}(u_1) \geq u_1a'_{21}(u_1) \geq 0$  and  $4a_{12}(u_2) \geq u_2a'_{12}(u_2) \geq 0$  for all  $u_1, u_2 \geq 0$ , then  $H_\varepsilon A_\varepsilon$  is positive definite and for all  $z = (z_1, z_2)^\top \in \mathbb{R}^2$ ,*

$$z^\top H_\varepsilon A_\varepsilon z \geq z^\top H A z + \varepsilon \left( \frac{\alpha_{10} + a_{11}(u_1)}{u_1} z_1^2 + \frac{\alpha_{20} + a_{22}(u_2)}{u_2} z_2^2 \right).$$

If  $p_1$  and  $p_2$  are given by (15),  $HA$  is positive definite if  $s \geq 1$ ,  $\alpha_{11}\alpha_{22} \geq (1 - 1/s)\alpha_{12}\alpha_{21}$ , and  $H_\varepsilon A_\varepsilon$  is positive definite if additionally  $s \leq 4$ . The proof also works in the case  $s < 1$ ; in this situation the restriction  $\alpha_{11}\alpha_{22} \geq (1 - 1/s)\alpha_{12}\alpha_{21}$  is not needed.



*Proof.* We have

$$HA = \begin{pmatrix} (p_1/u_1 + a'_{11})a'_{21} & a'_{12}a'_{21} \\ a'_{12}a'_{21} & (p_2/u_2 + a'_{22})a'_{12} \end{pmatrix}.$$

Then, for  $z = (z_1, z_2)^\top \in \mathbb{R}^2$ ,

$$\begin{aligned} z^\top HAz &= \frac{1}{s} a'_{12}a'_{21} \left( \sqrt{\frac{u_2}{u_1}} z_1 + \sqrt{\frac{u_1}{u_2}} z_2 \right)^2 + \left( \frac{\alpha_{10} + a_{11}}{u_1} + \frac{1}{u_1} \left( a_{12} - \frac{1}{s} u_2 a'_{12} \right) \right) a'_{21} z_1^2 \\ &\quad + \left( \frac{\alpha_{20} + a_{22}}{u_2} + \frac{1}{u_2} \left( a_{21} - \frac{1}{s} u_1 a'_{21} \right) \right) a'_{12} z_2^2 \\ &\quad + 2 \left( 1 - \frac{1}{s} \right) a'_{21} a'_{12} z_1 z_2 + a'_{11} a'_{21} z_1^2 + a'_{12} a'_{22} z_2^2. \end{aligned}$$

The last three terms are nonnegative for all  $z_1, z_2 \in \mathbb{R}$  if and only if  $a'_{11}a'_{21} \geq 0$  and  $a'_{12}a'_{22} \geq (1 - 1/s)a'_{21}a'_{12}$ . This shows the first statement of the lemma.

Next, we compute

$$\begin{aligned} H_\varepsilon A_\varepsilon &= HA + \varepsilon \begin{pmatrix} 1/u_1 & 0 \\ 0 & 1/u_2 \end{pmatrix} A(u) + \varepsilon H \begin{pmatrix} u_2 & 0 \\ 0 & u_1 \end{pmatrix} + \varepsilon^2 \begin{pmatrix} u_2/u_1 & 0 \\ 0 & u_1/u_2 \end{pmatrix} \\ &= HA + \varepsilon \begin{pmatrix} p_1/u_1 + a'_{11} + a'_{21}(u_2/u_1) & a'_{12} \\ a'_{21} & p_2/u_2 + a'_{22} + a'_{12}(u_1/u_2) \end{pmatrix} \\ &\quad + \varepsilon^2 \begin{pmatrix} u_2/u_1 & 0 \\ 0 & u_1/u_2 \end{pmatrix}. \end{aligned}$$

The first and the third matrix on the right-hand side are positive definite. Therefore, we need to analyze only the second matrix, called  $M$ :

$$\begin{aligned} z^\top Mz &= \left( \frac{p_1}{u_1} + a'_{11} + a'_{21} \frac{u_2}{u_1} \right) z_1^2 + \left( \frac{p_2}{u_2} + a'_{22} + a'_{12} \frac{u_1}{u_2} \right) z_2^2 + (a'_{21} + a'_{12}) z_1 z_2 \\ &= \frac{1}{2} a'_{21} \left( \sqrt{\frac{2u_2}{u_1}} z_1 + \sqrt{\frac{u_1}{2u_2}} z_2 \right)^2 + \frac{1}{2} a'_{12} \left( \sqrt{\frac{u_2}{2u_1}} z_1 + \sqrt{\frac{2u_1}{u_2}} z_2 \right)^2 \\ &\quad + \left( \frac{\alpha_{10} + a_{11} + a_{12}}{u_1} + a'_{11} + a'_{21} \frac{u_2}{u_1} - \left( a'_{21} + \frac{1}{4} a'_{12} \right) \frac{u_2}{u_1} \right) z_1^2 \\ &\quad + \left( \frac{\alpha_{20} + a_{21} + a_{22}}{u_2} + a'_{22} + a'_{12} \frac{u_1}{u_2} - \left( \frac{1}{4} a'_{21} + a'_{12} \right) \frac{u_1}{u_2} \right) z_2^2 \\ &\geq \left( \frac{\alpha_{10} + a_{11}}{u_1} + \frac{1}{u_1} \left( a_{12} - \frac{1}{4} u_2 a'_{12} \right) \right) z_1^2 + \left( \frac{\alpha_{20} + a_{22}}{u_2} + \frac{1}{u_2} \left( a_{21} - \frac{1}{4} u_1 a'_{21} \right) \right) z_2^2 \\ &\geq \frac{\alpha_{10} + a_{11}}{u_1} z_1^2 + \frac{\alpha_{20} + a_{22}}{u_2} z_2^2. \end{aligned}$$

This ends the proof.  $\square$

**Remark 12.** The positive semi-definiteness of  $HA$  can be proved by only assuming that  $\det A \geq 0$ . Indeed, setting  $HA = (c_{ij})$  and observing that  $HA$  is symmetric, for  $z = (z_1, z_2)^\top \in \mathbb{R}^2$ ,

$$z^\top HAz = c_{12} \left( \sqrt[4]{\frac{c_{11}}{c_{22}}} z_1 + \sqrt[4]{\frac{c_{22}}{c_{11}}} z_2 \right)^2 + \sqrt{\frac{c_{11}}{c_{22}}} (\sqrt{c_{11}c_{22}} - c_{12}) z_1^2$$

$$\begin{aligned}
& + \sqrt{\frac{c_{22}}{c_{11}}} (\sqrt{c_{11}c_{22}} - c_{12}) z_2^2 \\
& \geq (\sqrt{c_{11}c_{22}} - c_{12}) \min \left\{ \sqrt{\frac{c_{11}}{c_{22}}}, \sqrt{\frac{c_{22}}{c_{11}}} \right\} |z|^2,
\end{aligned}$$

and positive semi-definiteness follows if  $c_{11} > 0$ ,  $c_{22} > 0$ , and  $c_{11}c_{22} - c_{12}^2 = (\sqrt{c_{11}c_{22}} + c_{12})(\sqrt{c_{11}c_{22}} - c_{12}) \geq 0$ . Now,  $c_{11} \geq \alpha_{10}a'_{21}(u_1)/u_1 > 0$  for all  $u_1 > 0$ ,  $c_{22} > 0$  for all  $u_2 > 0$ , and, setting  $\partial_i p_j = \partial p_j / \partial u_i$ ,

$$c_{11}c_{22} - c_{12}^2 = \frac{a'_{21}a'_{12}}{u_1u_2}(p_1 + u_1\partial_1 p_1)(p_2 + u_2\partial_2 p_2) - (a'_{21}a'_{21})^2 = \frac{a'_{21}a'_{21}}{u_1u_2} \det A(u),$$

showing the claim.  $\square$

*Step 2: Solution of an approximate problem.* Set  $w = (Dh)^{-1}(u)$ . Then  $u = u(w)$ . The matrix  $B_\varepsilon(w) = A_\varepsilon(u(w))H_\varepsilon^{-1}(u(w))$  writes as

$$B_\varepsilon = \begin{pmatrix} u_1 p_1 / a'_1 + u_1^2 a'_1 / a'_{21} + \varepsilon u_1 u_2 / a'_{21} & u_1 u_2 \\ u_1 u_2 & u_2 p_2 / a'_2 + u_2^2 a'_{22} / a'_2 + \varepsilon u_1 u_2 / a'_2 \end{pmatrix}.$$

By Lemma 11,  $B_\varepsilon$  is positive semi-definite since  $z^\top B_\varepsilon z = (H_\varepsilon^{-1}z)^\top (H_\varepsilon A_\varepsilon)(H_\varepsilon^{-1}z) \geq 0$ . In view of Lemma 5, there exists a weak solution  $w^k$  to the approximate problem (18) satisfying the discrete entropy inequality (19) with  $C_f = 0$ . Summing this inequality from  $j = 1, \dots, k$  and employing the identity  $\nabla w^k : B_\varepsilon(w^k) \nabla w^k = \nabla u^k : H_\varepsilon A_\varepsilon \nabla u^k$  and Lemma 11, we find that

$$\begin{aligned}
& \int_\Omega h(u^k) dx + \tau \sum_{j=1}^k \int_\Omega \left( \frac{\alpha_{10} + a_{11}(u_1^j)}{u_1^j} a'_{21}(u_1^j) |\nabla u_1^j|^2 + \frac{\alpha_{20} + a_{22}(u_2^j)}{u_2^j} a'_{12}(u_2^j) |\nabla u_2^j|^2 \right) dx \\
& + \varepsilon \tau \sum_{j=1}^k \int_\Omega \left( \frac{\alpha_{10} + a_{11}(u_1^j)}{u_1^j} |\nabla u_1^j|^2 + \frac{\alpha_{20} + a_{22}(u_2^j)}{u_2^j} |\nabla u_2^j|^2 \right) dx \\
& + \varepsilon \tau \int_\Omega \left( \sum_{|\alpha|=m} |D^\alpha w|^2 + |w|^2 \right) dx \leq \int_\Omega h(u^0) dx.
\end{aligned}$$

*Step 3: Uniform estimates.* The growth conditions (48) and the above entropy estimate imply that

$$\begin{aligned}
& \tau \sum_{j=1}^k \sum_{i=1}^2 \int_\Omega ((u_i^j)^{s-2} + (u_i^j)^{2(s-1)}) |\nabla u_i^j|^2 dx \leq C, \\
& \varepsilon \tau \sum_{j=1}^k \sum_{i=1}^2 \int_\Omega ((u_i^j)^{-1} + (u_i^j)^{s-1}) |\nabla u_i^j|^2 dx \leq C.
\end{aligned}$$

Furthermore, since (48) shows that  $h(u) \geq C(u_1^s + u_2^s - 1)$  for some  $C > 0$ , the estimate

$$\int_\Omega ((u_1^k)^s + (u_2^k)^s) dx \leq \int_\Omega h(u^k) dx + C \leq C$$

yields a uniform bound for  $u_i^k$  in  $L^\infty(0, T; L^s(\Omega))$ . Note that by the Poincaré inequality,

$$\|(u_i^k)^s\|_{L^2(0, T; L^2(\Omega))} \leq C \|\nabla (u_i^k)^s\|_{L^2(0, T; L^2(\Omega))} + C \|(u_i^k)^s\|_{L^2(0, T; L^1(\Omega))} \leq C,$$

which implies that  $((u_i^k)^s)$  is bounded in  $L^2(0, T; H^1(\Omega))$ . Thus, defining  $u_i^{(\tau)}$  as in the proof of Theorem 2, we have proved the following uniform estimates:

$$(51) \quad \sum_{i=1}^2 \|u_i^{(\tau)}\|_{L^\infty(0, T; L^s(\Omega))} \leq C,$$

$$(52) \quad \sum_{i=1}^2 (\|(u_i^{(\tau)})^{s/2}\|_{L^2(0, T; H^1(\Omega))} + \|(u_i^{(\tau)})^s\|_{L^2(0, T; H^1(\Omega))}) \leq C,$$

$$(53) \quad \sqrt{\varepsilon} \sum_{i=1}^2 (\|(u_i^{(\tau)})^{1/2}\|_{L^2(0, T; H^1(\Omega))} + \|w_i^{(\tau)}\|_{L^2(0, T; H^m(\Omega))}) \leq C.$$

We also need an estimate for the time derivative. Let  $r = 2s(d+1)/(d(\sigma+1)) > 1$ ,  $1/r + 1/r' = 1$ , and  $\phi \in L^{r'}(0, T; X)$ , where  $X = \{\phi \in W^{m, \infty}(\Omega) : \nabla \phi \cdot \nu = 0 \text{ on } \partial\Omega\}$ . Then, observing that  $\operatorname{div}(A(u^{(\tau)}) \nabla u^{(\tau)})_i = \Delta(u_i^{(\tau)} p_i(u_i^{(\tau)}))$ , an integration by parts gives

$$\begin{aligned} \frac{1}{\tau} \left| \int_0^T \int_\Omega (u_i^{(\tau)} - \sigma_\tau u_i^{(\tau)}) \phi dx dt \right| &\leq \left| \int_0^T \int_\Omega u_i^{(\tau)} p_i(u_i^{(\tau)}) \Delta \phi dx dt \right| \\ &\quad + \varepsilon \int_0^T \|w_i^{(\tau)}\|_{H^m(\Omega)} \|\phi\|_{H^m(\Omega)} dt \\ &\leq \|u_i^{(\tau)} p_i(u_i^{(\tau)})\|_{L^r(0, T; L^1(\Omega))} \|\phi\|_{L^{r'}(0, T; W^{2, \infty}(\Omega))} \\ &\quad + \varepsilon \|w_i^{(\tau)}\|_{L^2(0, T; H^m(\Omega))} \|\phi\|_{L^2(0, T; H^m(\Omega))}. \end{aligned}$$

We estimate the norm of  $u_i^{(\tau)} p_i(u_i^{(\tau)})$ . We infer from the Gagliardo-Nirenberg inequality that  $L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^1(\Omega)) \hookrightarrow L^{2+2/d}(\Omega_T)$  and thus, by (51) and (52),  $((u_i^{(\tau)})^s)$  is bounded in  $L^{2+2/d}(\Omega_T)$ , where  $\Omega_T = \Omega \times (0, T)$ . The growth condition on  $p_i$  gives the estimate

$$\begin{aligned} \|u_i^{(\tau)} p_i(u_i^{(\tau)})\|_{L^r(\Omega_T)} &\leq C \left( 1 + \sum_{i=1}^2 \|u_i^{(\tau)}\|_{L^{(\sigma+1)r}(\Omega_T)}^{\sigma+1} \right) \\ (54) \quad &= C \left( 1 + \sum_{i=1}^2 \|u_i^{(\tau)}\|_{L^{2s(d+1)/d}(\Omega_T)}^{\sigma+1} \right) \leq C. \end{aligned}$$

We conclude that

$$\frac{1}{\tau} \left| \int_\tau^T \int_\Omega (u_i^{(\tau)} - \sigma_\tau u_i^{(\tau)}) \phi dx dt \right| \leq C \|\phi\|_{L^2(0, T; H^m(\Omega))}$$

and

$$(55) \quad \tau^{-1} \|u^{(\tau)} - \sigma_\tau u^{(\tau)}\|_{L^r(\tau, T; X')} \leq C.$$

*Step 3: The limit  $(\tau, \varepsilon) \rightarrow 0$ .* By estimates (52), (53), and (55), there exists a subsequence which is not relabeled such that, as  $(\tau, \varepsilon) \rightarrow 0$ ,

$$\begin{aligned} (u_i^{(\tau)})^s &\rightharpoonup y \quad \text{weakly in } L^2(0, T; H^1(\Omega)), \\ \tau^{-1} (u^{(\tau)} - \sigma_\tau u^{(\tau)}) &\rightharpoonup \partial_t u \quad \text{weakly in } L^r(0, T; X'), \\ \varepsilon w^{(\tau)} &\rightarrow 0 \quad \text{strongly in } L^2(0, T; H^m(\Omega)). \end{aligned}$$

Estimates (52) and (55) allow us also to apply the Aubin lemma in the version of [18], such that, for a subsequence,  $u^{(\tau)} \rightarrow u$  strongly in  $L^{2s}(0, T; L^{2q}(\Omega))$  as  $(\tau, \varepsilon) \rightarrow 0$ , where  $q \geq 2$  is such that  $H^1(\Omega) \hookrightarrow L^q(\Omega)$ . Because of the uniform bound of  $(u_i^{(\tau)})$  in  $L^{2s(d+1)/d}(\Omega_T)$  and the a.e. convergence of  $(u_i^{(\tau)})$ , we find that, up to a subsequence,  $u_i^{(\tau)} \rightarrow u_i$  strongly in  $L^p(\Omega_T)$  for all  $p < 2s(d+1)/d$  and in particular for  $p = 2s$ . By (52) and the above strong convergence, again up to a subsequence,  $u_i p_i(u^{(\tau)}) \rightarrow u_i p_i(u)$  a.e. in  $\Omega_T$ . Because of the bound (54), the dominated convergence theorem shows that  $u_i^{(\tau)} p_i(u^{(\tau)}) \rightarrow u_i p_i(u)$  strongly in  $L^r(\Omega_T)$ ,

Now, let  $\phi = (\phi_1, \phi_2) \in L^{r'}(0, T; X)^2$ . After an integration by parts,  $u^{(\tau)}$  solves

$$\begin{aligned} & \int_0^T \langle \tau^{-1}(\sigma_\tau u^{(\tau)} - u^{(\tau)}), \phi \rangle dt + \sum_{i=1}^2 \int_0^T \int_\Omega u_i^{(\tau)} p_i(u^{(\tau)}) \Delta \phi_i dx dt \\ & + \varepsilon \int_0^T \int_\Omega \left( \sum_{|\alpha|=m} D^\alpha w^{(\tau)} \cdot D^\alpha \phi + w^{(\tau)} \cdot \phi \right) dx dt = 0. \end{aligned}$$

The above convergence results are sufficient to pass to the limit  $(\tau, \varepsilon) \rightarrow 0$ , which yields

$$\int_0^T \langle \partial_t u, \phi \rangle dt + \sum_{i=1}^2 \int_0^T \int_\Omega u_i p_i(u) \Delta \phi_i dx dt = 0.$$

This proves the theorem.

## 6. FURTHER RESULTS AND OPEN PROBLEMS

In this section, we discuss some further results and open problems related to cross-diffusion systems and entropy methods.

- (i) *Dirichlet boundary conditions:* Dirichlet boundary conditions may be treated under some conditions on the nonlinearities. To see this, we assume that  $u_i = u_{D,i}$  on  $\partial\Omega$  for  $i = 1, \dots, n$  and we set  $w_D = Dh(u_D)$ . For simplicity, we assume that  $u_D = (u_{D,1}, \dots, u_{D,n})$  depends on the spatial variable only. Then, using the test function  $w - w_D$  in (1), we obtain

$$\begin{aligned} & \frac{d}{dt} \int_\Omega (h(u) - w_D u) dx + \int_\Omega \nabla u : (D^2 h)(u) A(u) \nabla u dx \\ (56) \quad & = \int_\Omega \nabla w_D : A(u) \nabla u dx + \int_\Omega f(u) \cdot Dh(u) dx - \int_\Omega f(u) \cdot w_D dx. \end{aligned}$$

If  $h(u)$  grows superlinearly (such as  $u_i \log u_i$ ) and  $w_D$  is smooth, we may estimate the first integral on the left-hand side as  $\int_\Omega (h(u) - w_D u) dx \geq \frac{1}{2} \int_\Omega h(u) dx - C$  for some constant  $C > 0$ . Under Hypotheses H2' and H3, the second integral on the left-hand side and the second integral on the right-hand side are estimated as follows:

$$\begin{aligned} \int_\Omega \nabla u : (D^2 h)(u) A(u) \nabla u dx & \geq C \sum_{i=1}^n \int_\Omega |\nabla \tilde{\alpha}_i(u_i)|^2 dx, \\ \int_\Omega f(u) \cdot Dh(u) dx & \leq C_f \int_\Omega (1 + h(u)) dx. \end{aligned}$$

Then, assuming that  $A(u)$  and  $f(u)$  are such that, for instance,

$$|A(u)\nabla u| \leq \sum_{i=1}^n |\nabla \tilde{\alpha}_i(u_i)|, \quad |f(u)| \leq C(1 + h(u)),$$

we can estimate the remaining integrals on the right-hand side of (56). In a similar way, we may treat Robin boundary conditions by combining the ideas for Dirichlet and homogeneous Neumann conditions.

- (ii) *Uniqueness of weak solutions:* It is possible to prove the uniqueness of weak solutions to (1)-(2) (with  $f = 0$ ) under the assumption that  $A(u)$  can be written as a gradient, i.e.  $A(u) = \nabla \Phi(u)$  for some monotone function  $\Phi : D \rightarrow \mathbb{R}^n$ , by employing the  $H^{-1}$  method. Unfortunately, the monotonicity assumption is rather strong since it requires that  $\nabla \Phi(u) = A(u)$  is positive semi-definite.

Another uniqueness result can be obtained if we are able to write the diffusion equation in terms of the entropy variable (see (3)), for instance in the approximated setting. For this result, we need to suppose that  $B(w) = \nabla \Phi(w)$  and that  $\Phi \circ Dh$  is monotone. Let  $w^{(1)}$  and  $w^{(2)}$  be two weak solutions to (2)-(3) (with  $f = 0$ ) with the same initial data. Furthermore, let  $v \in L^2(0, T; H^1(\Omega; \mathbb{R}^n))$  be the weak solution to  $-\Delta v = u(w^{(1)}) - u(w^{(2)})$  in  $\Omega$  and  $\nabla v \cdot \nu = 0$  on  $\partial\Omega$ . Then

$$\begin{aligned} \frac{1}{2} \int_{\Omega} |\nabla v(t)|^2 dx &= \int_0^T \langle \partial_t (u(w^{(1)}) - u(w^{(2)})), v \rangle dt \\ &= - \int_0^T \int_{\Omega} \nabla (\Phi(w^{(1)}) - \Phi(w^{(2)})) : \nabla v dx dt \\ &= - \int_0^T \int_{\Omega} (\Phi(w^{(1)}) - \Phi(w^{(2)})) \cdot (u(w^{(1)}) - u(w^{(2)})) dx dt \\ &= - \int_0^T \int_{\Omega} (\Phi(Dh(u^{(1)})) - \Phi(Dh(u^{(2)}))) \cdot (u^{(1)} - u^{(2)}) dx dt \leq 0, \end{aligned}$$

by the monotonicity of  $\Phi \circ Dh$ , where  $u^{(i)} = u(w^{(i)})$  for  $i = 1, 2$ . This shows that  $v = 0$  and hence  $u(w^{(1)}) = u(w^{(2)})$ . Uniqueness results under weaker conditions are an open problem (at least to our best knowledge). In fact, it is well known that this question is delicate since non-uniqueness is well known even for some scalar equations; see, e.g., [31].

- (iii) *Quadratic reaction terms:* In this paper, our focus was rather on the diffusion part than on the reaction part. The question is whether the results can be extended to diffusion systems with, say, quadratic reaction rates as they arise in reversible chemistry. For instance, a global existence result for the following problem is generally unknown:

$$\partial_t u - \operatorname{div}(A(u)\nabla) = f(u) \quad \text{in } \Omega, \quad f_i(u) = (-1)^i (u_1 u_3 - u_2 u_4),$$

for  $i = 1, 2, 3, 4$ , with boundary and initial conditions (2). If  $A(u)$  is the diagonal matrix  $A = \operatorname{diag}(d_1, \dots, d_n)$  with smooth functions  $d_i = d_i(x, t) \geq 0$ , the global existence of weak solutions was shown in [22]. This result is based on the observation that  $h(u) = \sum_{i=1}^4 u_i (\log u_i - 1)$  possesses an  $L^2$  estimate. This is proved by using the duality method of M. Pierre [49], and the proof is based on the diagonal structure of  $A$ . A global existence result for a system modeling the more general reaction  $A_1 + A_2 \rightleftharpoons A_3 + A_4$  was given in [10] in the two-dimensional case, the three-dimensional case being an open

problem. Reaction-diffusion systems with diffusivities  $d_i$  depending on  $u$  and quadratic rate functions were recently analyzed in [6], but still only in the diagonal case.

- (iv) *Long-time behavior of solutions:* We expect that the long-time behavior of weak solutions to (1)-(2) with  $f = 0$  can be proven under some additional conditions. For specific diffusion matrices, the exponential decay has been already shown; see [33, 34]. The idea is to estimate the entropy dissipation in terms of the entropy. For instance, let  $u_\infty = (u_{\infty,1}, \dots, u_{\infty,n})$  be a constant steady-state to (1)-(2). To simplify, we assume that the entropy is given by

$$\mathcal{H}^* = \sum_{i=1}^n \int_{\Omega} u_i \log \frac{u_i}{u_\infty} dx,$$

and that the entropy dissipation can be estimated from below according to

$$(57) \quad \int_{\Omega} \nabla w : B(w) \nabla w dx = \int_{\Omega} \nabla u : (D^2 h) A(u) \nabla u dx \geq \lambda \mathcal{H}^*$$

for some  $\lambda > 0$ . Then the entropy-dissipation inequality (4) becomes

$$\frac{d\mathcal{H}}{dt} + \lambda \mathcal{H} \leq 0,$$

and Gronwall's lemma implies exponential convergence in terms of the relative entropy  $\mathcal{H}^*$ . By the Csiszár-Kullback inequality, we conclude the exponential decay in the  $L^1$  norm,  $\|u(t) - u_\infty\|_{L^1(\Omega)} \leq C \exp(-\lambda t/2)$  for  $t > 0$ . For details, we refer to [11].

The main task is to derive the entropy-dissipation relation (57). According to Hypothesis H2', we need to prove

$$\sum_{i=1}^n \int_{\Omega} |\nabla \tilde{\alpha}_i(u_i)|^2 dx \geq \lambda \sum_{i=1}^n \int_{\Omega} u_i \log \frac{u_i}{u_\infty} dx.$$

If  $\tilde{\alpha}_i(u_i) = u_i^{1/2}$  (for instance), this inequality is a consequence of the logarithmic Sobolev inequality. For more general functions  $\tilde{\alpha}_i$  and power-type entropy densities, one may employ Beckner-type inequalities; see [12, Lemma 2]. The general case, however, is an open problem.

- (v) *Entropies for population models:* It is an open problem to find an entropy for the general population model with diffusion matrix  $A = (a_{ij})$  defined in (67). Just a simple combination of the entropies for the systems analyzed in Theorems 3 and 4, i.e. summing (12) and (16), seems to be not sufficient. It is also an open problem if an entropy for the population system with matrix (13), where  $p_1$  and  $p_2$  are given by (15), exists without restrictions on the coefficients  $\alpha_{ij}$  (except positivity).
- (vi) *Gradient systems and geodesic convexity:* There might to be a relation between our entropy formulation and the gradient structure for reaction-diffusion systems developed by Liero and Mielke [43] but no results are known so far. Using purely differential methods, they have shown the geodesic  $\lambda$ -convexity for particular reaction-cross-diffusion systems. It is an open problem if such geometric properties also hold for the examples presented in Section 2.
- (vii) *Coupled PDE-ODE problems:* Coupled reaction-diffusion-ODE problems occur, for instance, in chemotaxis-haptotaxis systems modeling cancer invasion [57]. Denoting by  $u_1$ ,  $u_2$ ,  $u_3$  the densities of the cancer cells, matrix-degrading enzymes, extracellular matrix,

respectively, the evolution is governed by (1) with the exemplary diffusion matrix

$$A(u) = \begin{pmatrix} a_{11}(u) & a_{12}(u) & a_{13}(u) \\ a_{21}(u) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It is an open problem whether these systems possess an entropy structure and whether our method can be extended to such problems. One idea could be to reduce the problem to the subsystem  $(u_1, u_2)$  and to estimate terms involving  $\nabla u_3$  directly by differentiating the ODE for  $u_3$ .

- (viii) *Bounded weak solutions:* A challenging task is to determine all diffusion systems whose weak solutions are bounded. As a first step in this direction, we consider  $n = 2$  and diffusion matrices  $A(u) = (a_{ij})$  depending linearly on the variables,

$$a_{ij}(u) = \alpha_{ij} + \beta_{ij}u_1 + \gamma_{ij}u_2, \quad i, j = 1, 2,$$

where  $\alpha_{ij}, \beta_{ij}, \gamma_{ij} \in \mathbb{R}$ . We wish to find conditions on the coefficients for which the logarithmic entropy density (6) satisfies Hypothesis H2. Requiring that  $B = A(D^2h)^{-1}$  is symmetric, we can fix some coefficients,

$$\begin{aligned} \alpha_{12} &= \alpha_{21} = \gamma_{12} = \beta_{21} = 0, & \beta_{22} &= \beta_{11} - \gamma_{21}, \\ \gamma_{11} &= \gamma_{22} - \beta_{12}, & \gamma_{21} &= \beta_{12} + \alpha_{22} - \alpha_{11}. \end{aligned}$$

The matrix  $(c_{ij}) = (D^2h)A$  is positive semi-definite if and only if

$$\begin{aligned} (58) \quad c_{11} &= \frac{1}{u_1(1-u_1-u_2)} \left( (\beta_{12} - \gamma_{22})u_2^2 + (-\beta_{11} + \beta_{12} + \alpha_{22} - \alpha_{11})u_1u_2 \right. \\ &\quad \left. + \beta_{11}u_1 + (-\alpha_{11} - \beta_{12} + \gamma_{22})u_2 + \alpha_{11} \right) \geq 0, \\ (59) \quad \det(c_{ij}) &= \frac{1}{u_1u_2(1-u_1-u_2)} \left( \beta_{11}(\beta_{11} - \beta_{12} - \alpha_{22} + \alpha_{11})u_1^2 - \gamma_{22}(\beta_{12} - \gamma_{22})u_2^2 \right. \\ &\quad \left. + (\alpha_{11}\gamma_{22} - \alpha_{22}\gamma_{22} - \beta_{11}\beta_{12} + 2\beta_{11}\gamma_{22} - \beta_{12}\gamma_{22})u_1u_2 \right. \\ &\quad \left. + (\alpha_{11}^2 - \alpha_{11}\alpha_{22} + \alpha_{11}\beta_{11} - \alpha_{11}\beta_{12} + \alpha_{22}\beta_{11})u_1 \right. \\ &\quad \left. + (\alpha_{11}\gamma_{22} - \alpha_{22}\beta_{12} + \alpha_{22}\gamma_{22})u_2 + \alpha_{11}\alpha_{22} \right) \geq 0 \end{aligned}$$

holds for all  $u_1, u_2 > 0$  such that  $u_1 + u_2 < 1$ , where admissible ranges for the remaining five parameters  $\alpha_{11}, \alpha_{22}, \beta_{11}, \beta_{12}, \gamma_{22}$  have to be determined. Examples are (i) the volume-filling model of Burger et al. (see Section 2.1) and (ii) the model defined by (9), where

- (i)  $\alpha_{11} = 1, \alpha_{22} = \beta, \beta_{11} = 0, \beta_{12} = 1, \gamma_{22} = 0,$
- (ii)  $\alpha_{11} = 1, \alpha_{22} = 1, \beta_{11} = -1, \beta_{12} = -1, \gamma_{22} = -1.$

Other examples can be easily constructed. For instance, for  $\alpha_{11} = \alpha_{22} = \beta_{11} = \beta_{12} = 1$  we have

$$A(u) = \begin{pmatrix} 1 + u_1 + (c_{22} - 1)u_2 & u_1 \\ u_2 & 1 + c_{22}u_2 \end{pmatrix},$$

and  $(c_{ij}) = (D^2h)A$  is positive semi-definite if and only if

$$\begin{aligned} c_{11} &= \frac{(1-u_2)(1+(c_{22}-1)u_2)+u_1}{u_1(1-u_1-u_2)} \geq 0 \quad \text{and} \\ \det(c_{ij}) &= (1+(c_{22}-1)u_2)(1+u_1+c_{22}u_2) \geq 0, \end{aligned}$$

which is the case if  $0 \leq c_{22} < \infty$ . However, it seems to be difficult to solve inequalities (58)-(59) in the general situation. Possibly, techniques from quadratic optimization with inequality constraints and quantifier elimination may help.

#### APPENDIX A. RELATIONS TO NON-EQUILIBRIUM THERMODYNAMICS

We show that the entropy variable  $w = Dh(u)$ , defined in Section 1, is strongly related to the chemical potentials of a fluid mixture and that the particular change of unknowns associated with the logarithmic entropy density (6) is related to a special choice of the thermodynamic activities.

We introduce first the thermodynamic setting. Consider a fluid consisting of  $N$  components with the same molar mass under isobaric and isothermal conditions. We write  $\rho_i$  instead of  $u_i$  to denote the mass density of the  $i$ th component. The evolution of the mass densities is governed by the mass balance equations

$$(60) \quad \partial_t \rho_i + \operatorname{div} J_i = 0, \quad i = 1, \dots, N,$$

where  $J_i$  are the diffusion fluxes. We have assumed that the barycentric velocity vanishes and that there are no chemical reactions. Furthermore, we assume for simplicity that the total mass density is constant,  $\sum_{j=1}^N \rho_j = 1$ .

Let  $s(\rho) = s(\rho_1, \dots, \rho_N)$  be the thermodynamic entropy of the system. Then the chemical potentials  $\mu_i$  are defined (in the isothermal case) by

$$\mu_i = -\frac{\partial s}{\partial \rho_i}, \quad i = 1, \dots, N,$$

where here and in the following we set physical constants (like temperature) equal to one. Neglecting also body forces and the (irreversible) stress tensor, the diffusion fluxes  $J_i$  can be written as (see [21, Chapter IV, (15)] or [5, Formula (170)])

$$(61) \quad J_i = -\sum_{j=1}^{N-1} L_{ij} \nabla(\mu_j - \mu_N), \quad i = 1, \dots, N,$$

where  $L_{ij}$  are some diffusion coefficients such that  $(L_{ij})$  is positive definite. Once an explicit expression for the thermodynamic entropy is determined, equations (60)-(61) are closed.

Now, we explain the relation of the entropy variables to the above setting. Since  $\rho_N = 1 - \sum_{j=1}^{N-1} \rho_j$ , we may express the density of the last component in terms of the others such that we can introduce

$$h(\rho_1, \dots, \rho_{N-1}) := -s\left(\rho_1, \dots, \rho_{N-1}, 1 - \sum_{j=1}^{N-1} \rho_j\right),$$

and in fact,  $h$  corresponds to the (mathematical) entropy introduced in Section 1. With this notation, the entropy variables become

$$w_i = \frac{\partial h}{\partial \rho_i} = -\frac{\partial s}{\partial \rho_i} + \frac{\partial s}{\partial \rho_N} = -(\mu_i - \mu_N), \quad i = 1, \dots, N-1,$$

which relates the entropy variables to the chemical potentials. Moreover, comparing the flux vector  $J = B(w) \nabla w$  from (3) with (61), we see that the diffusion matrix  $B(w)$  coincides with  $(L_{ij})$ .



For the second statement, we recall that the chemical potentials of a mixture of ideal gases can be formulated as

$$(62) \quad \mu_i = \mu_i^0 + \log \rho_i, \quad i = 1, \dots, N,$$

where  $\mu_i^0$  is the Gibbs energy which generally depends on temperature and pressure. Since we have supposed an isobaric, isothermal situation,  $\mu_i^0$  is constant and, for simplicity, we set  $\mu_i^0 = 0$  for  $i = 1, \dots, N$ . In order to model non-ideal gases, it is usual in thermodynamics to introduce the thermodynamic activity  $a_i$  and the activity coefficient  $\gamma_i$  by

$$\mu_i = \mu_i^0 + \log a_i = \log a_i, \quad \text{where } a_i = \gamma_i \rho_i.$$

If  $\gamma_i = 1$ , we recover the ideal-gas case. In the volume-filling case, Fuhrmann [28] has chosen  $\gamma_i = 1 + \sum_{j=1}^{N-1} a_j$  for numerical purposes. Then

$$\rho_i = \frac{a_i}{\gamma_i} = \frac{a_i}{1 + \sum_{j=1}^{N-1} a_j},$$

and since  $a_i = \exp(\mu_i)$ , it follows that

$$\rho_i = \frac{e^{\mu_i}}{1 + \sum_{j=1}^{N-1} e^{\mu_j}}, \quad i = 1, \dots, N-1.$$

which corresponds to the inverse transformation (8) if we identify  $\mu_i$  with  $w_i$ . This expression can be derived directly from (62). Indeed, if  $\mu_i^0 = 0$ , we obtain  $w_i = -(\mu_i - \mu_N) = -\log(\rho_i/\rho_N)$  with  $\rho_N = 1 - \sum_{j=1}^{N-1} \rho_j$ , and inverting these relations, we find that

$$\rho_i = \frac{e^{w_i}}{1 + \sum_{j=1}^{N-1} e^{w_j}}, \quad i = 1, \dots, N-1.$$

## APPENDIX B. FORMAL DERIVATION OF TWO-SPECIES POPULATION MODELS

We derive formally cross-diffusion systems from a master equation for a continuous-time, discrete-space random walk in the macroscopic limit. We consider only random walks on one-dimensional lattices but the derivation extends to higher dimensions in a straightforward way. The lattice is given by cells  $x_i$  ( $i \in \mathbb{Z}$ ) with the uniform cell distance  $h > 0$ . The densities of the populations in the  $i$ th cell at time  $t > 0$  are denoted by  $u_1(x_i, t)$  and  $u_2(x_i, t)$ , respectively. We assume that the population species  $u_1$  and  $u_2$  move from the  $i$ th cell into the neighboring  $(i \pm 1)$ th cells with transition rates  $S_i^\pm$  and  $T_i^\pm$ , respectively. Then the master equations can be formulated as follows [47]:

$$(63) \quad \partial_t u_1(x_i) = S_{i-1}^+ u_1(x_{i-1}) + S_{i+1}^- u_1(x_{i+1}) - (S_i^+ + S_i^-) u_1(x_i),$$

$$(64) \quad \partial_t u_2(x_i) = T_{i-1}^+ u_2(x_{i-1}) + T_{i+1}^- u_2(x_{i+1}) - (T_i^+ + T_i^-) u_2(x_i),$$

where  $i \in \mathbb{Z}$ . We further suppose that the transition rates  $S_i^\pm$  and  $T_i^\pm$  depend on the departure cell  $i$  and the arrival cells  $i \pm 1$ :

$$S_i^\pm = \sigma_0 p_1(u_1(x_i), u_2(x_i)) q_1(1 - u_1(x_{i\pm 1}) - u_2(x_{i\pm 1})),$$

$$T_i^\pm = \sigma_0 p_2(u_1(x_i), u_2(x_i)) q_2(1 - u_1(x_{i\pm 1}) - u_2(x_{i\pm 1})),$$

where  $\sigma_0 > 0$  is some number. Abbreviating  $p_j(x_i) = p_j(u_1(x_i), u_2(x_i))$  and  $q_j(x_i) = q_j(1 - u_1(x_i) - u_2(x_i))$ , the master equations (63)-(64) become

$$\sigma_0^{-1} \partial_t u_1(x_i) = p_1(x_{i-1}) q_1(x_i) u_1(x_{i-1}) + p_1(x_{i+1}) q_1(x_i) u_1(x_{i+1})$$

$$\begin{aligned}
(65) \quad & -p_1(x_i)(q_1(x_{i-1}) + q_1(x_{i+1}))u_1(x_i), \\
& \sigma_0^{-1}\partial_t u_2(x_i) = p_2(x_{i-1})q_2(x_i)u_2(x_{i-1}) + p_2(x_{i+1})q_2(x_i)u_2(x_{i+1}) \\
(66) \quad & -p_2(x_i)(q_2(x_{i-1}) + q_2(x_{i+1}))u_2(x_i).
\end{aligned}$$

The functions  $p_1(x_i)$  and  $p_2(x_i)$  model the tendency of the species to leave the cell  $i$ , whereas  $q_1(x_i)$  and  $q_2(x_i)$  describe the probability to move into the cell  $i$ . The latter functions allow us to model the so-called volume-filling effect. Indeed, we may interpret  $u_1$  and  $u_2$  as volume fractions satisfying  $u_1 + u_2 \leq 1$ . Then  $1 - u_1 - u_2$  describes the volume fraction not occupied by the two species. If the  $i$ th cell is fully occupied, i.e.  $u_1(x_i) + u_2(x_i) = 1$ , and  $q_j(1 - u_1(x_i) - u_2(x_i)) = q_j(0) = 0$ , the probability to move into the  $i$ th cell is zero.

In order to derive a macroscopic model, we perform a formal Taylor expansion of  $u_j(x_{i\pm 1})$ ,  $p_j(x_{i\pm 1})$ , and  $q_j(x_{i\pm 1})$  around  $u_j(x_i)$  up to second order. Furthermore, we assume a diffusive scaling, i.e.  $\sigma_0 = 1/h^2$ . Substituting the expansions into (65) and (66) and performing the formal limit  $h \rightarrow 0$ , it follows that (see [47] for details)

$$\begin{aligned}
\partial_t u_1 &= \partial_x(a_{11}(u)\partial_x u_1 + a_{12}(u)\partial_x u_2), \\
\partial_t u_2 &= \partial_x(a_{21}(u)\partial_x u_1 + a_{22}(u)\partial_x u_2),
\end{aligned}$$

where  $u = (u_1, u_2)$ . The diffusion coefficients are given by

$$\begin{aligned}
(67) \quad & a_{11}(u) = p_1(u)q_1(u_3) + u_1(\partial_1 p_1(u)q_1(u_3) + p_1(u)q_1'(u_3)), \\
& a_{12}(u) = u_1(\partial_2 p_1(u)q_1(u_3) + p_1(u)q_1'(u_3)), \\
& a_{21}(u) = u_2(\partial_1 p_2(u)q_2(u_3) + p_2(u)q_2'(u_3)), \\
& a_{22}(u) = p_2(u)q_2(u_3) + u_2(\partial_2 p_2(u)q_2(u_3) + p_2(u)q_2'(u_3)),
\end{aligned}$$

where  $\partial_i p_j = \partial p_j / \partial u_i$  and  $u_3 = 1 - u_1 - u_2$ . In several space dimensions, the argumentation is the same but the computations are more involved. We obtain equation (1) with  $f = 0$  and the diffusion matrix  $A$  with coefficients  $a_{ij}$  as above.

It seems very difficult—if not impossible—to explore the entropy structure of this equation in full generality. Therefore, we investigated two special cases in this paper. First, we assumed that  $p_1 = p_2 = 1$ ,  $q_1 = q$ , and  $q_2 = \beta q$ , where  $\beta > 0$ . This corresponds to the volume-filling case (10). Second, we have set  $q_1 = q_2 = 1$ . This gives the matrix (13).

#### APPENDIX C. A VARIANT OF THE AUBIN COMPACTNESS LEMMA

**Lemma 13.** *Let  $(y_\tau)$ ,  $(z_\tau)$  be sequences which are piecewise constant in time with step size  $\tau > 0$  and which are bounded in  $L^\infty(0, T; L^\infty(\Omega))$ . Let  $(y_\tau)$  be relatively compact in  $L^2(0, T; L^2(\Omega))$ , i.e., up to subsequences which are not relabeled,  $y_\tau \rightarrow y$  strongly in  $L^2(0, T; L^2(\Omega))$  and  $z_\tau \rightharpoonup^* z$  weakly\* in  $L^\infty(0, T; L^\infty(\Omega))$  as  $\tau \rightarrow 0$ . Finally, let*

$$\begin{aligned}
& \|y_\tau\|_{L^2(0, T; H^1(\Omega))} \leq C, \\
& \|y_\tau z_\tau\|_{L^2(0, T; H^1(\Omega))} + \tau^{-1} \|\pi_\tau z_\tau - z_\tau\|_{L^2(\tau, T; (H^1(\Omega))')} \leq C,
\end{aligned}$$

where  $(\pi_\tau z_\tau)(\cdot, t) = z_\tau(\cdot, t + \tau)$  for  $0 < t \leq T - \tau$ . Then there exists a subsequence (not relabeled) such that  $y_\tau z_\tau \rightarrow yz$  strongly in  $L^p(0, T; L^p(\Omega))$  for all  $p < \infty$ .

Note that the result would follow from the Aubin compactness lemma if  $y_\tau$  was bounded from below by a positive constant, since in this situation, it would suffice to apply the Aubin lemma in the version of [25] to infer the strong convergence of (a subsequence of)  $(z_\tau)$  in  $L^2(0, T; L^2(\Omega))$  which, together with the strong convergence of  $(y_\tau)$ , would give the result.

*Proof.* The proof is inspired from [8, Section 4.4] but parts of the proof are different. The idea is to prove that

$$\lim_{(h,k) \rightarrow 0} \int_{\Omega_T} ((y_\tau z_\tau)(x+h, t+k) - (y_\tau z_\tau)(x, t))^2 dx dt = 0 \quad \text{uniformly in } \tau > 0,$$

where  $\Omega_T = \Omega \times (0, T)$  and  $y_\tau(\cdot, t)$ ,  $z_\tau(\cdot, t)$  are extended by zero for  $T \leq t \leq T+k$ . Then the result follows from the lemma of Kolmogorov-Riesz [7, Theorem 4.26] and the  $L^\infty$  boundedness of  $y_\tau z_\tau$ . We write

$$\begin{aligned} & \int_{\Omega_T} ((y_\tau z_\tau)(x+h, t+k) - (y_\tau z_\tau)(x, t))^2 d(x, t) \\ & \leq 2 \int_{\Omega_T} ((y_\tau z_\tau)(x+h, t+k) - (y_\tau z_\tau)(x, t+k))^2 d(x, t) \\ & \quad + 2 \int_{\Omega_T} ((y_\tau z_\tau)(x, t+k) - (y_\tau z_\tau)(x, t))^2 d(x, t) = I_1 + I_2. \end{aligned}$$

For the estimate of  $I_1$ , we integrate over the line segment  $[x, x+h]$  and employ a standard extension operator in  $L^2$ :

$$I_1 \leq \int_{\Omega_T} \int_0^1 h^2 |\nabla(y_\tau z_\tau)(x+sh, t+k)|^2 ds d(x, t) \leq Ch^2 \|\nabla(y_\tau z_\tau)\|_{L^2(\Omega_T)} \leq Ch^2,$$

where here and in the following,  $C > 0$  denotes a generic constant.

For the second integral, we have

$$\begin{aligned} I_2 & \leq 4 \int_{\Omega_T} (y_\tau(x, t+k) - y_\tau(x, t))^2 z_\tau(x, t+k)^2 d(x, t) \\ & \quad + 4 \int_{\Omega_T} y_\tau(x, t)^2 (z_\tau(x, t+k) - z_\tau(x, t))^2 d(x, t) = I_{21} + I_{22}. \end{aligned}$$

Since  $(z_\tau)$  is bounded in  $L^\infty(0, T; L^\infty(\Omega))$ , we can estimate as follows:

$$(68) \quad I_{21} \leq C \int_{\Omega_T} (y_\tau(x, t+k) - y_\tau(x, t))^2 d(x, t).$$

By assumption,  $(y_\tau)$  is relatively compact in  $L^2(\Omega_T)$ . By the inverse of the lemma of Kolmogorov-Riesz [7, Exercise 4.34], the right-hand side of (68) converges to zero as  $k \rightarrow 0$  uniformly in  $\tau > 0$ . Furthermore,

$$\begin{aligned} I_{22} & = \int_{\Omega_T} y_\tau(x, t)^2 z_\tau(x, t)(z_\tau(x, t) - z_\tau(x, t+k)) d(x, t) \\ & \quad + \int_{\Omega_T} y_\tau(x, t+k)^2 z_\tau(x, t+k)(z_\tau(x, t+k) - z_\tau(x, t)) d(x, t) \\ & \quad + \int_{\Omega_T} (y_\tau(x, t)^2 - y_\tau(x, t+k)^2) z_\tau(x, t+k)(z_\tau(x, t+k) - z_\tau(x, t)) d(x, t) \\ & = J_1 + J_2 + J_3. \end{aligned}$$

Using Lemma 5 in [18] and the bounds on  $(y_\tau)$ ,  $(z_\tau)$ , the first integral can be estimated as

$$\begin{aligned} J_1 & \leq \|y_\tau^2 z_\tau\|_{L^2(0, T; H^1(\Omega))} \|\pi_k z_\tau - z_\tau\|_{L^2(0, T-k; H^1(\Omega)')} \\ & \leq Ck^{1/2} \tau^{-1} \|\pi_\tau z_\tau - z_\tau\|_{L^1(0, T-\tau; H^1(\Omega)')} \leq Ck^{1/2}, \end{aligned}$$

and this converges to zero as  $k \rightarrow 0$  uniformly in  $\tau > 0$ . The same conclusion holds for  $J_2$ . Finally, in view of the  $L^\infty$  boundedness of  $(y_\tau)$  and  $(z_\tau)$ , the third integral becomes

$$\begin{aligned} J_3 &\leq \int_{\Omega_T} (y_\tau(x, t+k) + y_\tau(x, t))(y_\tau(x, t+k) - y_\tau(x, t))z_\tau(x, t+k) \\ &\quad \times (z_\tau(x, t+k) - z_\tau(x, t))d(x, t) \\ &\leq C \int_{\Omega_T} |y_\tau(x, t+k) - y_\tau(x, t)|d(x, t). \end{aligned}$$

Because of the relative compactness of  $(y_\tau)$  in  $L^2$ , the right-hand side converges to zero uniformly in  $\tau > 0$ . This finishes the proof.  $\square$

## REFERENCES

- [1] N. Alikakos.  $L^p$  bounds of solutions of reaction-diffusion equations. *Commun. Part. Diff. Eqs.* 4 (1979), 827-868.
- [2] H. Amann. Dynamic theory of quasilinear parabolic systems. III. Global existence. *Math. Z.* 202 (1989), 219-250.
- [3] H. Amann. Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems. In: H.J. Schmeisser and H. Triebel (editors), *Function Spaces, Differential Operators and Nonlinear Analysis*, pages 9-126. Teubner, Stuttgart, 1993.
- [4] D. Bothe. On the Maxwell-Stefan equations to multicomponent diffusion. In: *Progress in Nonlinear Differential Equations and their Applications*, pp. 81-93. Springer, Basel, 2011.
- [5] D. Bothe and W. Dreyer. Continuum thermodynamics of chemically reacting fluid mixtures. Preprint, 2014. [arXiv:1401.5991](#).
- [6] D. Bothe and G. Rolland. Global existence for a class of reaction-diffusion systems with mass action kinetics and concentration-dependent diffusivities. Preprint, 2014. [arXiv:1402.4846](#).
- [7] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York, 2011.
- [8] M. Burger, M. Di Francesco, J.-F. Pietschmann, and B. Schlake. Nonlinear cross-diffusion with size exclusion. *SIAM J. Math. Anal.* 42 (2010), 2842-2871.
- [9] M. Burger, B. Schlake, and W.-T. Wolfram. Nonlinear Poisson-Nernst-Planck equations for ion flux through confined geometries. *Nonlinearity* 25 (2012), 961-990.
- [10] J. A. Cañizo, L. Desvillettes, and K. Fellner. Improved duality estimates and applications to reaction-diffusion equations. Preprint, 2013. [arXiv:1304.4040](#).
- [11] J. A. Carrillo, A. Jüngel, P. Markowich, G. Toscani, and A. Unterreiter. Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Monatsh. Math.* 133 (2001), 1-82.
- [12] C. Chainais-Hillairet, A. Jüngel, and S. Schuchnigg. Entropy-dissipative discretization of nonlinear diffusion equations and discrete Beckner inequalities. Preprint, 2013. [arXiv:1303.3791](#).
- [13] G. Chavent and J. Jaffre. *Mathematical Models and Finite Elements for Reservoir Simulation*. North Holland, Amsterdam, 1986.
- [14] L. Chen and A. Jüngel. Analysis of a multi-dimensional parabolic population model with strong cross-diffusion. *SIAM J. Math. Anal.* 36 (2004), 301-322.
- [15] L. Chen and A. Jüngel. Analysis of a parabolic cross-diffusion population model without self-diffusion. *J. Diff. Eqs.* 224 (2006), 39-59.
- [16] L. Chen and A. Jüngel. Analysis of a parabolic cross-diffusion semiconductor model with electron-hole scattering. *Commun. Part. Diff. Eqs.* 32 (2007), 127-148.
- [17] X. Chen and A. Jüngel. Analysis of an incompressible Navier-Stokes-Maxwell-Stefan system. Submitted for publication, 2013. [arXiv:1310.3376](#).
- [18] X. Chen, A. Jüngel, and J.-G. Liu. A note on Aubin-Lions-Dubinskii lemmas. To appear in *Acta Appl. Math.*, 2014. [arXiv:1305.6235](#).
- [19] P. Degond, S. Génieys, and A. Jüngel. Symmetrization and entropy inequality for general diffusion equations. *C. R. Acad. Sci. Paris* 325 (1997), 963-968.

- [20] P. Degond, S. Génieys, and A. Jüngel. A system of parabolic equations in nonequilibrium thermodynamics including thermal and electrical effects. *J. Math. Pures Appl.* 76 (1997), 991-1015.
- [21] S. de Groot and P. Mazur. *Non-equilibrium Thermodynamics*. Dover, New York, 1962.
- [22] L. Desvillettes, K. Fellner, M. Pierre, and J. Vovelle. Global existence for quadratic systems of reaction-diffusion. *Adv. Nonlin. Stud.* 7 (2007), 491-511.
- [23] L. Desvillettes, T. Lepoutre, and A. Moussa. Entropy, duality and cross diffusion. Preprint, 2013. **arXiv:1302.1028**.
- [24] L. Desvillettes and A. Trescases. New results for triangular reaction cross-diffusion system. Preprint CMLA 2013-13, ENS Cachan, France, 2013.
- [25] M. Dreher and A. Jüngel. Compact families of piecewise constant functions in  $L^p(0, T; B)$ . *Nonlin. Anal.* 75 (2012), 3072-3077.
- [26] F. Duzaar, G. Mingione, and K. Steffen. Parabolic systems with polynomial growth and regularity. *Mem. Amer. Math. Soc.* 214 (2011), 118 pages.
- [27] L. Evans. *Partial Differential Equations*. Amer. Math. Soc., Providence, 1998.
- [28] J. Fuhrmann. Personal communication, 2014.
- [29] G. Galiano, M. Garzón, and A. Jüngel. Semi-discretization in time and numerical convergence of solutions of a nonlinear cross-diffusion population model. *Numer. Math.* 93 (2003), 655-673.
- [30] D. Gillespie, W. Nonner, and R. Eisenberg. Coupling Poisson-Nernst-Planck and density functional theory to calculate ion flux. *J. Phys. Condens. Matter* 14 (2002), 12129-12145.
- [31] A. Haraux and F. Weissler. Non-uniqueness for a semilinear initial value problem. *Indiana Univ. Math. J.* 31 (1982), 167-189.
- [32] T. Jackson and H. Byrne. A mechanical model of tumor encapsulation and transcapsular spread. *Math. Biosci.* 180 (2002), 307-328.
- [33] A. Jüngel and I. Stelzer. Entropy structure of a cross-diffusion tumor-growth model. *Math. Models Meth. Appl. Sci.* 22 (2012), 1250009 (26 pages).
- [34] A. Jüngel and I. Stelzer. Analysis of Maxwell-Stefan systems for multicomponent mixtures. To appear in *SIAM J. Math. Anal.*, 2014. **arXiv:1211.2394**.
- [35] S. Kawashima and Y. Shizuta. On the normal form of the symmetric hyperbolic-parabolic systems associated with the conservation laws. *Tohoku Math. J., II. Ser.* 40 (1988), 449-464.
- [36] J. Kim. Smooth solutions to a quasi-linear system of diffusion equations for a certain population model. *Nonlin. Anal.* 8 (1984), 1121-1144.
- [37] K. Küfner. Invariant regions for quasilinear reaction-diffusion systems and applications to a two population model. *NoDEA Nonlin. Diff. Eqs. Appl.* 3 (1996), 421-444.
- [38] O.A. Ladyženskaya, V.A. Solonnikov, and N.N. Ural'ceva. *Linear and Quasilinear Equations of Parabolic Type*. Amer. Math. Soc., Providence, 1988.
- [39] D. Le. Global existence for a class of strongly coupled parabolic systems. *Ann. Math.* 185 (2006), 133-154.
- [40] D. Le. Cross diffusion systems in  $n$  spatial dimensional domains. *Indiana Univ. Math. J.* 51 (2002), 625-643.
- [41] D. Le and T. Nguyen. Everywhere regularity of solutions to a class of strongly coupled degenerate parabolic systems. *Commun. Part. Diff. Eqs.* 31 (2006), 307-324.
- [42] T. Lepoutre, M. Pierre, and G. Rolland. Global well-posedness of a conservative relaxed cross diffusion system. *SIAM J. Math. Anal.* 44 (2012), 1674-1693.
- [43] M. Liero and A. Mielke. Gradient structures and geodesic convexity for reaction-diffusion systems. *Phil. Trans. Royal Soc. A.* 371 (2005), 20120346 (28 pages), 2013.
- [44] A. Lunardi. *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Birkhäuser, Basel, 1995.
- [45] M. Marion and R. Temam. Global existence for fully nonlinear reaction-diffusion systems describing multicomponent reactive flows. Preprint, 2013. **arXiv:1310.2624**.
- [46] P. Markowich. *The Stationary Semiconductor Device Equations*. Springer, Wien, 1986.
- [47] S. Ostrander. Macroscopic cross-diffusion models derived from spatially discrete continuous time microscopic models. *SIAM Undergrad. Res. Online* 4 (2011), <http://dx.doi.org/10.1137/10S010818>.
- [48] K. Painter and T. Hillen. Volume-filling and quorum-sensing in models for chemosensitive movement. *Canadian Appl. Math. Q.* 10 (2002), 501-543.
- [49] M. Pierre and D. Schmitt. Blowup in reaction-diffusion systems with dissipation of mass. *SIAM J. Math. Anal.* 28 (1997), 259-269.

- [50] J. Prüss. Maximal regularity for evolution equations in  $L^p$ -spaces. *Conf. Sem. Mat. Univ. Bari* 285 (2003), 1-39.
- [51] K. Redlinger. Invariant sets for strongly coupled reaction-diffusion systems under general boundary conditions. *Arch. Ration. Mech. Anal.* 108 (1989), 281-291.
- [52] D. Reznik. Generalised drift-diffusion model of bipolar transport in semiconductors. In: H. Ryssel and P. Pichler (eds.), *Simulation of Semiconductor Devices and Processes* 6 (1995), 254-257.
- [53] T. Roubíček. *Nonlinear Partial Differential Equations with Applications*. Birkhäuser, Basel, 2005.
- [54] N. Shigesada, K. Kawasaki, and E. Teramoto. Spatial segregation of interacting species. *J. Theor. Biol.* 79 (1979), 83-99.
- [55] S.-A. Shim. Uniform boundedness and convergence of solutions to cross-diffusion systems. *J. Diff. Eqs.* 185 (2002), 281-305.
- [56] J. Stará and O. John. Some (new) counterexamples of parabolic systems. *Comment. Math. Univ. Carolin.* 36 (1995), 503-510.
- [57] Y. Tao and C. Cui. A density-dependent chemotaxis-haptotaxis system modeling cancer invasion. *J. Math. Anal. Appl.* 367 (2010), 612-624.
- [58] R. Temam. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*. Second edition. Springer, New York, 1997.
- [59] A.M. Turing. The chemical basis of morphogenesis. *Phil. Trans. Roy. Soc. London, Ser. B* 237 (1952), 37-72.
- [60] D. Wrzosek. Volume filling effect in modelling chemotaxis. *Math. Model. Nat. Phenom.* 5 (2010), 123-147.
- [61] A. Yagi. Global solution to some quasilinear parabolic systems in population dynamics. *Nonlin. Anal.* 21 (1993), 603-630.

INSTITUTE FOR ANALYSIS AND SCIENTIFIC COMPUTING, VIENNA UNIVERSITY OF TECHNOLOGY, WIEDNER  
HAUPTSTRASSE 8-10, 1040 WIEN, AUSTRIA

*E-mail address:* juengel@tuwien.ac.at